THE WORLD ATLAS OF LANGUAGE STRUCTURES

EDITED BY
MARTIN HASPELMATH
MATTHEW S. DRYER
DAVID GIL
AND
BERNARD COMRIE

Oxford
LINGUISTICS

The World Atlas of Language Structures

This page intentionally left blank

The World Atlas of Language Structures

Edited by

Martin Haspelmath, Matthew S. Dryer,
David Gil, and Bernard Comrie

With the collaboration of

Hans-Jörg Bibiko, Hagen Jung, and Claudia Schmidt

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© M Haspelmath, M Dryer, D Gil, B Comrie 2005
The moral rights of the authors have been asserted
Database right Oxford University Press (marker)
First published 2005
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer
British Library Cataloguing in Publication Data
(Data available)
Library of Congress Cataloguing in Publication Data
(Data available)
ISBN 0 19 92559 1 (Hbk) 9780199255917
10 9 8 7 6 5 4 3 2 1
Typeset by Graphicraft Limited, Hong Kong
Printed in China
on acid-free paper by

Contents

Abbreviations ix
Acknowledgements xi
List of Contributors xiii

INTRODUCTION 1
Bernard Comrie, Matthew S. Dryer, David Gil, and Martin Haspelmath

TEXTS AND MAPS

A. Phonology

1 Consonant Inventories 10
Ian Maddieson
2 Vowel Quality Inventories 14
Ian Maddieson
3 Consonant–Vowel Ratio 18
Ian Maddieson
4 Voicing in Plosives and Fricatives 22
Ian Maddieson
5 Voicing and Gaps in Plosive Systems 26
Ian Maddieson
6 Uvular Consonants 30
Ian Maddieson
7 Glottalized Consonants 34
Ian Maddieson
8 Lateral Consonants 38
Ian Maddieson
9 The Velar Nasal (*ŋ*) 42
Gregory D. S. Anderson
10 Vowel Nasalization 46
John Hajek
11 Front Rounded Vowels 50
Ian Maddieson
12 Syllable Structure 54
Ian Maddieson
13 Tone 58
Ian Maddieson
14 Fixed Stress Locations 62
Rob Goedemans and Harry van der Hulst
15 Weight-Sensitive Stress 66
Rob Goedemans and Harry van der Hulst
16 Weight Factors in Weight-Sensitive Stress Systems 70
Rob Goedemans and Harry van der Hulst
17 Rhythm Types 74
Rob Goedemans and Harry van der Hulst
18 Absence of Common Consonants 78
Ian Maddieson
19 Presence of Uncommon Consonants 82
Ian Maddieson

B. Morphology

20 Fusion of Selected Inflectional Formatives 86
Balthasar Bickel and Johanna Nichols
21 Exponence of Selected Inflectional Formatives 90
Balthasar Bickel and Johanna Nichols
22 Inflectional Synthesis of the Verb 94
Balthasar Bickel and Johanna Nichols
23 Locus of Marking in the Clause 98
Johanna Nichols and Balthasar Bickel
24 Locus of Marking in Possessive Noun Phrases 102
Johanna Nichols and Balthasar Bickel
25 Locus of Marking: Whole-language Typology 106
Johanna Nichols and Balthasar Bickel
26 Prefixing versus Suffixing in Inflectional Morphology 110
Matthew S. Dryer
27 Reduplication 114
Carl Rubino
28 Case Syncretism 118
Matthew Baerman and Dunstan Brown
29 Syncretism in Verbal Person/Number Marking 122
Matthew Baerman and Dunstan Brown

C. Nominal Categories

Gender and Number

30 Number of Genders 126
Greville G. Corbett
31 Sex-based and Non-sex-based Gender Systems 130
Greville G. Corbett
32 Systems of Gender Assignment 134
Greville G. Corbett
33 Coding of Nominal Plurality 138
Matthew S. Dryer
34 Occurrence of Nominal Plurality 142
Martin Haspelmath
35 Plurality in Independent Personal Pronouns 146
Michael Daniel
36 The Associative Plural 150
Michael Daniel and Edith Moravcsik

Articles and Pronouns

37 Definite Articles 154
Matthew S. Dryer
38 Indefinite Articles 158
Matthew S. Dryer
39–40 Inclusive/Exclusive Forms for 'we' 162
Michael Cysouw
39 Inclusive/Exclusive Distinction in Independent Pronouns 162
40 Inclusive/Exclusive Distinction in Verbal Inflection 164
41 Distance Contrasts in Demonstratives 170
Holger Diessel
42 Pronominal and Adnominal Demonstratives 174
Holger Diessel
43 Third-Person Pronouns and Demonstratives 178
D. N. S. Bhat
44 Gender Distinctions in Independent Personal Pronouns 182
Anna Siewierska

45 Politeness Distinctions in Pronouns 186
Johannes Helmbrecht
46 Indefinite Pronouns 190
Martin Haspelmath
47 Intensifiers and Reflexive Pronouns 194
Ekkehard König and Peter Siemund (with Stephan Töpper)
48 Person Marking on Adpositions 198
Dik Bakker

Case

49 Number of Cases 202
Oliver A. Iggesen
50 Asymmetrical Case Marking 206
Oliver A. Iggesen
51 Position of Case Affixes 210
Matthew S. Dryer
52 Comitatives and Instrumentals 214
Thomas Stolz, Cornelia Stroh, and Aina Urdze

Numerals

53 Ordinal Numerals 218
Thomas Stolz and Ljuba N. Veselinova
54 Distributive Numerals 222
David Gil
55 Numeral Classifiers 226
David Gil
56 Conjunctions and Universal Quantifiers 230
David Gil

D. Nominal Syntax

57 Position of Pronominal Possessive Affixes 234
Matthew S. Dryer
58 Obligatory Possessive Inflection 238
Johanna Nichols and Balthasar Bickel
59 Possessive Classification 242
Johanna Nichols and Balthasar Bickel
60 Genitives, Adjectives, and Relative Clauses 246
David Gil
61 Adjectives without Nouns 250
David Gil
62 Action Nominal Constructions 254
Maria Koptjevskaja-Tamm
63 Noun Phrase Conjunction 258
Leon Stassen
64 Nominal and Verbal Conjunction 262
Martin Haspelmath

E. Verbal Categories

Tense and Aspect

65–68 Tense and Aspect 266
Östen Dahl and Viveka Velupillai
65 Perfective/Imperfective Aspect 267
66 The Past Tense 268
67 The Future Tense 270
68 The Perfect 272
69 Position of Tense-Aspect Affixes 282
Matthew S. Dryer

Modality

70 The Morphological Imperative 286
Johan van der Auwera and Ludo Lejeune (with Umarani Pappuswamy and Valentin Goussev)
71 The Prohibitive 290
Johan van der Auwera and Ludo Lejeune (with Valentin Goussev)
72 Imperative-Hortative Systems 294
Johan van der Auwera, Nina Dobrushina, and Valentin Goussev
73 The Optative 298
Nina Dobrushina, Johan van der Auwera, and Valentin Goussev
74 Situational Possibility 302
Johan van der Auwera and Andreas Ammann
75 Epistemic Possibility 306
Johan van der Auwera and Andreas Ammann
76 Overlap between Situational and Epistemic Modal Marking 310
Johan van der Auwera and Andreas Ammann
77 Semantic Distinctions of Evidentiality 314
Ferdinand de Haan
78 Coding of Evidentiality 318
Ferdinand de Haan

Suppletion

79 Suppletion According to Tense and Aspect 322
Ljuba N. Veselinova
80 Verbal Number and Suppletion 326
Ljuba N. Veselinova

F. Word Order

81 Order of Subject, Object, and Verb 330
Matthew S. Dryer
82 Order of Subject and Verb 334
Matthew S. Dryer
83 Order of Object and Verb 338
Matthew S. Dryer
84 Order of Object, Oblique, and Verb 342
Matthew S. Dryer (with Orin D. Gensler)
85 Order of Adposition and Noun Phrase 346
Matthew S. Dryer
86 Order of Genitive and Noun 350
Matthew S. Dryer
87 Order of Adjective and Noun 354
Matthew S. Dryer
88 Order of Demonstrative and Noun 358
Matthew S. Dryer
89 Order of Numeral and Noun 362
Matthew S. Dryer
90 Order of Relative Clause and Noun 366
Matthew S. Dryer
91 Order of Degree Word and Adjective 370
Matthew S. Dryer
92 Position of Polar Question Particles 374
Matthew S. Dryer

93 Position of Interrogative Phrases in Content Questions 378
Matthew S. Dryer

94 Order of Adverbial Subordinator and Clause 382
Matthew S. Dryer

95 Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase 386
Matthew S. Dryer

96 Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun 390
Matthew S. Dryer

97 Relationship between the Order of Object and Verb and the Order of Adjective and Noun 394
Matthew S. Dryer

G. Simple Clauses

98–99 Alignment of Case Marking 398
Bernard Comrie

98 Alignment of Case Marking of Full Noun Phrases 399

99 Alignment of Case Marking of Pronouns 400

100 Alignment of Verbal Person Marking 406
Anna Siewierska

101 Expression of Pronominal Subjects 410
Matthew S. Dryer

102 Verbal Person Marking 414
Anna Siewierska

103 Third-Person Zero of Verbal Person Marking 418
Anna Siewierska

104 Order of Person Markers on the Verb 422
Anna Siewierska

105 Ditransitive Constructions: The Verb 'Give' 426
Martin Haspelmath

Valence and Voice

106 Reciprocal Constructions 430
Elena Maslova and Vladimir P. Nedjalkov

107 Passive Constructions 434
Anna Siewierska

108 Antipassive Constructions 438
Maria Polinsky

109 Applicative Constructions 442
Maria Polinsky

110 Periphrastic Causative Constructions 446
Jae Jung Song

111 Nonperiphrastic Causative Constructions 450
Jae Jung Song

Negation and Questions

112 Negative Morphemes 454
Matthew S. Dryer

113 Symmetric and Asymmetric Standard Negation 458
Matti Miestamo

114 Subtypes of Asymmetric Standard Negation 462
Matti Miestamo

115 Negative Indefinite Pronouns and Predicate Negation 466
Martin Haspelmath

116 Polar Questions 470
Matthew S. Dryer

Predication

117 Predicative Possession 474
Leon Stassen

118 Predicative Adjectives 478
Leon Stassen

119 Nominal and Locational Predication 482
Leon Stassen

120 Zero Copula for Predicate Nominals 486
Leon Stassen

121 Comparative Constructions 490
Leon Stassen

H. Complex Sentences

122–123 Relativization Strategies 494
Bernard Comrie and Tania Kuteva

122 Relativization on Subjects 494

123 Relativization on Obliques 496

124 'Want' Complement Clauses 502
Martin Haspelmath

125 Purpose Clauses 506
Sonia Cristofaro

126 'When' Clauses 510
Sonia Cristofaro

127 Reason Clauses 514
Sonia Cristofaro

128 Utterance Complement Clauses 518
Sonia Cristofaro

I. Lexicon

129 Hand and Arm 522
Cecil H. Brown

130 Finger and Hand 526
Cecil H. Brown

131 Numeral Bases 530
Bernard Comrie

132–135 Colour Terms 534
Paul Kay and Luisa Maffi

132 Number of Nonderived Basic Colour Categories 534

133 Number of Basic Colour Categories 535

134 Green and Blue 535

135 Red and Yellow 536

136–137 Personal Pronouns 546
Johanna Nichols and David Peterson

136 M-T Pronouns 546

137 N-M Pronouns 546

138 Tea 554
Östen Dahl

J. Sign Languages

Sign Languages 558
Ulrike Zeshan

139 Irregular Negatives in Sign Languages 560
Ulrike Zeshan

140 Question Particles in Sign Languages 564
Ulrike Zeshan

K. Other

141 Writing Systems 568
Bernard Comrie

142 Paralinguistic Usages of Clicks 572
David Gil

REFERENCE MATERIAL

Genealogical Maps 576
Genealogical Language List 584
by Matthew S. Dryer
References 645
Language Name Index 664
WALS Code Index 685

Abbreviations

1	first person
2	second person
3	third person
A	agent-like argument
ABL	ablative
ABS	absolutive
ACC	accusative
ACCOMP	accompanied
ACR	actor
ACT	actual
ADEL	adelative
ADJ	adjective
ADVZ	adverbializer
AFF	affirmative
AGT	agent
ALL	allative
AN	action nominal
ANC	action nominal construction
ANIM	animate
ANTIP	antipassive
AOR	aorist
APPL	applicative
ART	article
AS	asseverative
ASP	aspect
ASSOC	associative
ASY	asymmetric
ATTR	attributive
AUD	auditory evidential
AUG	augmented
AUX	auxiliary
BEN	benefactive
C	common gender
CAUS	causative
CL	class (= noun class, gender)
CLF	classifier
CMPL	completive
CNTR	contrary to expectation marker
COLL	collective
COM	comitative
COMP	complementizer
COMPR	comparative
CONN	connective
CONNEG	connegative
CONSTR	construct
CONT	continuative, continuous
CONTEMP	contemporative
CONV	converb
COP	copula
CPW	categories per word
CRS	currently relevant state
DAT	dative
DECL	declarative
DEF	definite
DEG	degree word
DEM	demonstrative
DEP	dependent marker
DES	desire
DESID	desiderative
DIM	diminutive
DIR	direct
DIR.EVD	direct evidential
DIRL	directional
DIST.PST	distant past
DOBJ	direct object
DS	different subject
DU	dual
DUR	durative
EMPH	emphatic
EPENTH	epenthetic
EPV	expletive verbal suffix
ERG	ergative
EVD	evidential
EXCL	exclusive
F	feminine
FACT	factive
FAM	familiar
FIN	finite
FIN.AOR	finite aorist
FOC	focus
FUT	future
FV	verb-final vowel
GEN	genitive
HAB	habitual
HEST	hesternal past
HHON	super honorific
HOD	hodiernal past
HON	honorific
HORT	hortative
HUM	human
IE	Indo-European
ILL	illative
IMM.PRET	immediate preterite
IMM.PST	immediate past
IMP	imperative
IMPF	imperfective
IMPERS	impersonal
INAN	inanimate
INCEP	inceptive
INCOMPL	incompletive
IND	indicative
INDEF	indefinite
INDIR.EVD	indirect evidential
INFER	inferential evidential
INGR	ingressive
INSTR	instrumental
INTENT	intentional
INTER	interrogative
INTF	intensifier
INTGEN	intended genitive
INTR	intransitive
INV	inverse
IO	indirect object
IRR	irrealis
ITER	iterative
LIG	ligature
LOC	locative
LOCUT	locutor person marker
M	masculine
MEAS	measure word
MED	medial
MID	middle
MOD	modifier
N	neuter
NARR	narrative
NC	noun class
NEC	necessity
NEG	negation/negative
NHON	non-honorific
NMLZ	nominalizer/nominalization
NOM	nominative
NON.F	nonfeminine

NONFIN	nonfinite
NONFIN.AOR	nonfinite aorist
NONHUM	nonhuman
NONVIS	nonvisual
NP	noun phrase
NPST	non-past
NSG	nonsingular
NUM	numeral
O	object pronominal marker
OBJ	object
OBL	oblique
OBV	obviative
OPT	optative
P	patient-like argument
PASS	passive
PAT	patient
PATH	path locative
PCL	particle
PERF	perfect
PERS	personal
PFV	perfective
PHR.TERM	phrase terminal marker
PL	plural
PLUPERF	pluperfect
POS	possibility
POSS	possessive
POSTP	postposition
POT	potential
PP	prepositional/postpositional phrase
PRECONTEMP	precontemporal
PRED	predicative
PREF	prefix
PREP	preposition
PRES	present
PREV	preverb
PROG	progressive
PROL	prolative
PRON	pronoun
PROP	proper name
PROX	proximal/proximate
PRTV	partitive
PST	past
PST.CONT	past continuous
PST.PUNCT	past punctiliar
PSTBEFOREYEST	past before yesterday (= prehesternal)
PTCP	participle
PUNCT	punctual stem
PURP	purposive
Q	question-marker
QUOT	quotative
RDP	reduplication
REAL	realis
REC	recent (past)
RECP	reciprocal
REDUP	reduplication
REM.PST	remote past
REMOTE	remote
REPET	repetitive
RLZ	realized
RNR	result nominalizer
S	sole argument of the intransitive verb
SBJV	subjunctive
SENS	sensory evidential
SG	singular
SPEC	specific
SR	switch reference
SS	same subject
STAT	stative
SUBJ	subject
SUBORD	subordination
SUFF	suffix
SUP	superessive
SYM	symmetric
SYMASY	symmetric and asymmetric
T/A	tense/aspect
TD	time depth/proximality marker
TELIC	telic
TEMPRY	temporary
TH	thematic suffix
THM	theme (i.e. the semantic role)
TNS	tense
TOD.PST	today past
TOP	topic
TR	transitive
TRANSL	translative
TRI	trial
UNSP	unspecified
VBLZ	verbalizer
VENT	ventive
VIS	visual evidential
VP	verb phrase

Acknowledgements

An enterprise like *The World Atlas of Language Structures* is impossible without the help of more people than its editors and authors. We gratefully acknowledge their crucial contributions.

Foremost among our helpers were our collaborators in Leipzig (and Jakarta):

Claudia Schmidt, who was responsible for text uniformity and formatting, and spent many weeks on the consistency of the references (this part of her work is only visible in the electronic version);

Hagen Jung, who rescued the computational work on *WALS* from a crisis, and created the map files with competence, patience, and an excellent sense for beauty;

Hans-Jörg Bibiko, who programmed the interactive electronic version with great enthusiasm and creativity, proposing many features that had never occurred to the editors;

Orin Gensler, who copy-edited and proofread all the texts, corresponded with most of the authors and also helped eliminate some factual errors—we were lucky to have the assistance of the world's best linguistics copy-editor (not to mention his role as co-author);

and Brad Taylor, working at the Jakarta Field Station, who designed our working database and was always there when problems needed to be addressed.

Without the commitment and unfailing reliability of these people, *WALS* would have been much less than what it is now. Thank you so much, all of you!

We are also indebted to our student assistants in Leipzig, Martin Mehlberg, Tamar Khizanishvili, Luise Kempf, and Thomas Hanke, as well as the University of Leipzig computer scientists Christian Wolff, Georg Apitz, Marek Mahn, Christian Heinrich, and Gerrit Mattausch.

The Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology (Director Bernard Comrie) provided the funding for the editors between 1999 and 2005, funded several meetings of a subset of the authors, and allowed some of them to come to Leipzig for a more extended period to work on their contributions. We are especially grateful to Julia Cissewski, Claudia Büchel, and Peter Fröhlich for creating a perfect working environment.

This is also the place to acknowledge the help of the language experts who patiently answered many questions from the authors, their main reward being that the maps of the atlas now show many more dots for "their" language than would otherwise have been possible. The following 89 people generously agreed to serve as expert consultants:

Abkhaz	George Hewitt
Alamblak	Les Bruce
Amele	John Roberts
Apurinã	Sidi Facundes
Armenian (Eastern)	Jasmine Dum-Tragut
Awa Pit	Timothy Curnow
Bambara	Valentin Vydrine
Basque	José Hualde
Batak (Karo)	Geoff Woollams
Beja	Richard Hudson
Brahui	Josef Elfenbein
Chamorro	Sandra Chung
Chukchi	Michael Dunn
Coos (Hanis)	Matthew Dryer
Cree (Plains)	Rose-Marie Déchaine, Christoph Wolfart
Evenki	Igor Nedjalkov
Finnish	Pirkko Suihkonen
Georgian	George Hewitt
Gooniyandi	William McGregor
Greek (Modern)	Brian Joseph
Greenlandic (West)	Jeanette Sakel
Hausa	Mahamane L. Abdoulaye
Hebrew	David Gil
Hindi	Anvita Abbi
Hixkaryana	Desmond Derbyshire
Hungarian	István Kenesei
Hunzib	Helma van den Berg
Indonesian	David Gil
Ingush	Johanna Nichols
Japanese	Akiko Comrie
Kannada	S. N. Sridhar
Kayardild	Nicholas Evans
Ket	Edward Vajda
Khasi	Gracious (Grace) Temsen
Khmu'	Jan-Olof Svantesson
Kilivila	Gunter Senft
Kobon	John Davies
Korean	Ho-Min Sohn
Koromfe	John Rennison
Koyraboro Senni	Jeffrey Heath
Krongo	Mechthild Reh
Kutenai	Matthew Dryer
Lak	Konstantin Kazenin
Lakhota	Robert Van Valin
Lango	Michael Noonan
Latvian	Aleksandra Steinbergs
Lavukaleve	Angela Terrill
Lezgian	Martin Haspelmath
Makah	Matthew Davidson
Malagasy	Charles Randriamasimanana
Martuthunira	Alan Dench
Maybrat	Philomena Dol
Meithei	Shobhana Chelliah
Mixtec (Chalcatongo)	Monica Macaulay
Ndyuka	George Huttar, Mary Huttar
Nez Perce	Noel Rude
Ngiyambaa	Tamsin Donaldson
Nivkh	Elena Gruzdeva
Nunggubuyu	Jeffrey Heath
Oneida	Karin Michelson
Oromo (Harar)	Jonathan Owens
Paamese	Terry Crowley
Pirahã	Daniel Everett
Pitjantjatjara	Heather Bowe
Quechua (Imbabura)	Pieter Muysken
Rama	Colette Grinevald
Russian	Bernard Comrie
Semelai	Nicole Kruspe
Shipibo-Konibo	Pilar Valenzuela
Slave	Keren Rice
Taba	John Bowden
Tagalog	Carl Rubino
Trumai	Raquel Guirardello
Turkish	Jaklin Kornfilt
Vietnamese	Mark Alves
Wambaya	Rachel Nordlinger
Wari'	Dan Everett
Wichi	Kenneth Claesson
Wichita	David Rood
Yagua	Doris Payne, Tom Payne
Yidiny	R. M. W. Dixon
Yimas	William Foley
Yoruba	Ayo Bamgbose
Yukaghir (Kolyma)	Elena Maslova
Yup'ik (Central)	Anthony Woodbury

The research for Matthew Dryer's chapters was supported by Social Sciences and Humanities Research Council of Canada

Research Grants 410-810949, 410-830354, and 410-850540, and by National Science Foundation (United States) Research Grant BNS-9011190.

Matthew Dryer also wants to acknowledge a number of people who over the years worked on his typological database project as graduate student research assistants. These include, at the University of Alberta: Janice Adlington, Sandra Bellan, Somsak Boonsathorn, Diane Dennis, Randy Harris, Barbara Jacennik, Betty Karpinski, Murray Munro, Martha Smith, and Grace Wiebe; at UCLA: Mercedes Pérez; and at the University at Buffalo: Isao Honda, Anna Keusen (Margetts), Monica Madera, Heather Weber, and Byong-Seon Yang. Rob Malouf helped him programme his own database, which was the prototype of *WALS*.

Thanks also to anyone else whom we have inadvertently overlooked.

Needless to say, any errors and shortcomings should exclusively be blamed on the editors and authors. (The editors will keep a catalog of reported errors under <http://www.wals.info/>)

Martin Haspelmath, Matthew S. Dryer,
David Gil, Bernard Comrie

Leipzig, 23 July 2004

List of Contributors

Ammann, Andreas
Centre for Grammar, Cognition and Typology
University of Antwerp
Universiteitsplein 1
B-2610 Wilrijk
Belgium
andreas.ammann@ua.ac.be

Anderson, Gregory D. S.
8269 Jordan St. SE
Salem, OR 97301
USA
xakasboy@yahoo.com

Baerman, Matthew
Surrey Morphology Group
Linguistic, Cultural and Translation Studies
University of Surrey
GB-Guildford, Surrey GU2 5XH
United Kingdom
m.baerman@surrey.ac.uk

Bakker, Dik
Department of General Linguistics
University of Amsterdam
Spuistraat 210
NL-1012 VT Amsterdam
The Netherlands
d.bakker@uva.nl

Bhat, D. N. S.
402, 5th Cross
Roopanagar
Mysore 570 005
India
dnsbhat@sancharnet.in

Bickel, Balthasar
Institut für Linguistik
Universität Leipzig
Beethovenstr. 15
D-04107 Leipzig
Germany
bickel@uni-leipzig.de

Brown, Cecil H.
6039 Chandelle Circle
Pensacola, FL 32507
USA
chbrown@niu.edu

Brown, Dunstan
Surrey Morphology Group
Linguistic, Cultural and Translation Studies
University of Surrey
GB-Guildford, GU2 5XH
United Kingdom
d.brown@surrey.ac.uk

Comrie, Bernard
Max-Planck-Institut für evolutionäre Anthropologie
Deutscher Platz 6
D-04103 Leipzig
Germany
comrie@eva.mpg.de

Corbett, Greville G.
Surrey Morphology Group
Linguistic, Cultural and Translation Studies
University of Surrey
GB-Guildford, GU2 5XH
United Kingdom
g.corbett@surrey.ac.uk

Cristofaro, Sonia
Dipartimento di Linguistica
Università di Pavia
Corso Strada Nuova 65
I–27100 Pavia
Italy
russo@inrete.it

Cysouw, Michael
Max-Planck-Institut für evolutionäre Anthropologie
Deutscher Platz 6
D-04103 Leipzig
Germany
cysouw@eva.mpg.de

Dahl, Östen
Department of Linguistics
Stockholm University
S-106 91 Stockholm
Sweden
oesten@ling.su.se

Daniel, Michael
(Lomonosov Moscow State University, Institute of World Culture)
ul. Dovatora, dom 3 kv. 15
RF-119048 Moscow
Russia
daniel@qub.com

de Haan, Ferdinand
Department of Linguistics
University of Arizona
P.O. Box 210028
Tucson, AZ 85721
USA
fdehaan@u.arizona.de

Diessel, Holger
Institut für Anglistik/Amerikanistik
Universität Jena
Ernst-Abbe-Platz 8
D-07743 Jena
Germany
holger.diessel@uni-jena.de

Dobrushina, Nina R.
(State University Higher School of Economics, Department of Language and Literature)
ul. Dovatora, dom 3 kv. 15
RF-119048 Moscow
Russia
daniel@qub.com

Dryer, Matthew S.
Department of Linguistics
University at Buffalo
Buffalo, NY 14260
USA
dryer@buffalo.edu

Gensler, Orin D.
Max-Planck-Institut für evolutionäre Anthropologie
Deutscher Platz 6
D-04103 Leipzig
Germany
gensler@eva.mpg.de

Gil, David
Max-Planck-Institut für evolutionäre Anthropologie
Deutscher Platz 6
D-04103 Leipzig
Germany
gil@eva.mpg.de

Goedemans, Rob
Phonetics Lab
Universiteit Leiden
Cleveringaplaats 1
P.O. Box 9515
NL-2300 RA Leiden
The Netherlands
R.W.N.Goedemans@let.leidenuniv.nl

Goussev, Valentin
Institute of Linguistics of Russian Academy of Science
B. Kislovskij per., 1/12
RF-103009 Moscow
Russia
goussev@iling.msk.su

Hajek, John
Department of French, Italian and Spanish Studies
School of Languages
University of Melbourne
Parkville, Victoria 3052
Australia
j.hajek@unimelb.edu.au

Haspelmath, Martin
Max-Planck-Institut für evolutionäre Anthropologie
Deutscher Platz 6
D-04103 Leipzig
Germany
haspelmath@eva.mpg.de

Helmbrecht, Johannes
Seminar für Sprachwissenschaft
Universität Erfurt
Postfach 900221
D-99105 Erfurt
Germany
johannes.helmbrecht@uni-erfurt.de

Iggesen, Oliver A.
Rathenaustraße 14
44263 Dortmund
Germany
oliver_iggesen@yahoo.de

Kay, Paul
International Computer Science Institute
1947 Center Street
Berkeley, CA 94704
USA
kay@icsi.berkeley.edu

König, Ekkehard
Institut für Englische Philologie
Freie Universität Berlin
Gosslerstraße 2–4
D-14195 Berlin
Germany
koenig@zedat.fu-berlin.de

Koptjevskaja-Tamm, Maria
Department of Linguistics
Stockholm University
S-10691 Stockholm
Sweden
tamm@ling.su.se

Kuteva, Tania
Anglistik III
Heinrich-Heine-Universität Düsseldorf
Universitätsstr. 1
D-40223 Düsseldorf
Germany
kouteva@phil-fak.uni-duesseldorf.de

Lejeune, Ludo
Centre for Grammar, Cognition and Typology
University of Antwerp
Universiteitsplein 1
B-2610 Wilrijk
Belgium
(Ludo Lejeune passed away in 2003)

Maddieson, Ian
Department of Linguistics
University of California at Berkeley
1203 Dwinelle Hall
Berkeley, CA 94720-2650
USA
ianm@socrates.berkeley.edu

Maffi, Luisa
Terralingua
1766 Lanier Place NW
Washington, DC 20009
USA
maffi@terralingua.org

Maslova, Elena
2000 Walnut Ave., #J307
Fremont, CA 94538
USA
maslova@jps.net

Miestamo, Matti
General Linguistics
P.O. Box 9
FI-00014 University of Helsinki
Finland
matti.miestamo@ling.helsinki.fi

Moravcsik, Edith
Department of Linguistics
University of Wisconsin-Milwaukee
Milwaukee, WI 53201-0413
USA
edith@csd.uwm.edu

Nedjalkov, Vladimir P.
Institute of Linguistic Research
Tuchkov per., 9
RF-199053 St. Petersburg
Russia
nedjalkov@typology.spb.su

Nichols, Johanna
Department of Slavic Languages and Literatures
University of California at Berkeley
5416 Dwinelle Hall
Berkeley, CA 94720
USA
johanna@uclink.berkeley.edu

Pappuswamy, Umarani
Learning Research and Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260
USA
umarani@pitt.edu

Peterson, David A.
Linguistics & Cognitive Science
Dartmouth College
6086 Reed Hall
Hanover, NH 03755
USA
david.a.peterson@dartmouth.edu

Polinsky, Maria
Department of Linguistics
University of California at San Diego
La Jolla, CA 92093-0108
USA
polinsky@ling.ucsd.edu

Rubino, Carl
506 Dean Dr.
Rockville, MD 20851-1173
USA
crubi@tagalog.org

Siemund, Peter
Universität Hamburg
Institut für Anglistik und Amerikanistik
Von-Melle-Park 6
20146 Hamburg
Germany
peter.siemund@uni-hamburg.de

Siewierska, Anna
Department of Linguistics and Modern English Language
Lancaster University
Lancaster, LAI 4YT
United Kingdom
a.siewierska@lancaster.ac.uk

Song, Jae Jung
Linguistics Programme
School of Language, Literature, and Performing Arts
University of Otago
P.O. Box 56
Dunedin
New Zealand
jaejung.song@stonebow.otago.ac.nz

Stassen, Leon
Department of Linguistics
Radboud University Nijmegen
Erasmusplein 1
NL-6525 GG Nijmegen
The Netherlands
l.stassen@let.ru.nl

Stolz, Thomas
Universität Bremen
FB 10, Linguistik
Postfach 330 440
D-28334 Bremen
Germany
stolz@uni-bremen.de

Stroh, Cornelia
Universität Bremen
FB 10, Linguistik
Postfach 330 440
D-28334 Bremen
Germany
komtyp@uni-bremen.de

Töpper, Stephan
Institut für Englische Philologie
Freie Universität Berlin
Gosslerstraße 2–4
D-14195 Berlin
Germany
toepper@zedat.fu-berlin.de

Urdze, Aina
Universität Bremen
FB 10, Linguistik
Postfach 330 440
D-28334 Bremen
Germany
aina@uni-bremen.de

van der Auwera, Johan
Centre for Grammar, Cognition and Typology
University of Antwerp
Universiteitsplein 1
B-2610 Wilrijk
Belgium
johan.vanderauwera@ua.ac.be

van der Hulst, Harry
Department of Linguistics, U-1145
University of Connecticut
337 Mansfield Road
Storrs, CT 06269-1145
USA
vdhulst@uconnvm.uconn.edu

Velupillai, Viveka
Institut für Anglistik
Phil I, Haus B, 3. u. 4. Stock
Otto-Behaghel-Str. 10B
D-35394 Gießen
Germany
viveka.velupillai@email.de

Veselinova, Ljuba N.
Department of Linguistics
Stockholm University
S-10691 Stockholm
Sweden
ljuba@ling.su.se

Zeshan, Ulrike
Max-Planck-Institut für Psycholinguistik
Postbus 310
NL-6500 AH Nijmegen
The Netherlands
Ulrike.Zeshan@mpi.nl

This page intentionally left blank

Introduction

BERNARD COMRIE, MATTHEW S. DRYER,
DAVID GIL, AND MARTIN HASPELMATH

1 What and why?

The World Atlas of Language Structures (*WALS*) provides the reader with 142 maps showing the geographical distribution of structural linguistic features. It is a quite novel type of atlas. Linguists have long worked with maps showing the geographical distribution of languages (i.e. the areas where most of their speakers live), and a complete world atlas was published a decade ago (Moseley and Asher 1994; see also *Ethnologue*, Grimes 2000). For well over a century, linguists have also produced atlases that show the geographical distribution of linguistic features in the dialects of a language. *WALS* is the first feature atlas on a worldwide scale. It can be thought of as a kind of dialect atlas of the "dialects" of Human Language. But it differs from dialect atlases in an important way. While dialect atlases show the geography of substantive linguistic features (such as particular cognate sounds, or particular words), *WALS* shows only **structural features**, i.e. abstract features of the language system that can be compared across unrelated languages.

Linguists interested in **linguistic typology**—the systematic study of the ways in which the languages of the world vary structurally and of the limits to this variation—have recently begun to ask questions relating to the geographical distribution of different values for structural linguistic features. For instance, they may want to know whether languages with a particular word order in the clause, say subject–verb–object as in English (*the farmer killed the duckling*), are found only in one part of the world, whether they are distributed more or less evenly across different parts of the world, or whether some in-between scenario holds. Although previous work has been able to provide some answers to some questions of this kind, these answers have hitherto been by and large unsystematic, often reflecting more the intuitive feel that a particular linguist has for the geographical distribution of the feature in question rather than a consistent sampling of the world's languages in order to answer the question. *The World Atlas of Language Structures* aims to provide just this kind of systematic answer, since the authors of the individual chapters, each dealing with a particular linguistic feature, have set out to be as comprehensive (within their sampling limits) as possible in the mapping of variants for that feature across the languages of the world. The printed atlas provides a visual overview of this distribution, by using different coloured dots for the different feature values. In addition, the text that accompanies each printed map provides an explanation of the feature values and of their assignment, as well as discussion of patterns of geographical distribution and of the relevance of the chapter to theoretical issues. The interactive electronic version (on the accompanying CD-ROM) provides much more information, including the possibility of zooming in on particular geographical areas that seem of particular interest for a particular feature. It provides access to the original data and bibliographical or other sources that underlie the atlas's database. It enables the user to manipulate data, for instance to calculate what percentage of the world's languages have a particular feature value, or to see whether there is a correlation between particular values of different features.

Although the main users of *WALS* will be those interested in linguistic typology, the atlas is also relevant to the interests of other linguists. For instance, certain theoretical approaches have been criticized for being based too heavily on languages exhibiting particular geographically restricted feature values, and *WALS* will enable those interested in testing such criticisms to see whether they do indeed hold. Thus, much recent work on relative clauses has been based heavily on the kind of construction found in the major European literary languages, such as literary English *the man whom I saw*, and has been criticized for extending the analysis appropriate for such languages to other languages with radically different relative clause types. While the material on relative clauses in the atlas does not, of course, directly address the issue of whether a particular approach to the analysis of relative clauses is valid cross-linguistically, it does show that the distribution of the "European" type of relative clause (in the terminology used here: the relative pronoun strategy) is by and large restricted to Europe, thus at least calling into question theoretical approaches that rely on so geographically restricted a typological variant (see Chapters 122–123).

Both typologists and other linguists are interested in questions relating to **correlations among different features**, such as whether there is indeed a tendency for occurrence of prepositions (rather than postpositions) to correlate with verb–object order in the clause, and for postpositions to correlate with object–verb order; whether there is a tendency for noun–adjective order in the noun phrase to correlate with verb–object order in the clause, and adjective–noun order with object–verb order. The relevant *WALS* data, in particular the maps that can be generated in the electronic version by combining data from different individual maps, suggest that there is validity to the postulated correlation in the first case (adpositions and verb position), but not in the second (adjectives and verb position) (see Chapters 95–97).

Previous work in typology has provided extensive results regarding the ways in which languages vary structurally and regarding correlations among different features. As noted above, *WALS* provides significant further contributions to these areas, especially with the data provided by the electronic version. However, because of the maps, *WALS* provides an especially significant contribution to the field of **areal typology**, which seeks to establish whether particular geographical distributions are the result of language contact among neighbouring languages.

The maps vary with respect to the degree to which the features show areal patterns. On some maps, clear geographical patterns emerge. On Map 83, for example, one finds a clear distribution of the two orders of object and verb among the languages of Eurasia. In Europe, most of the languages are VO (placing the verb before the object), while to the east of this is a huge area covering much of Asia, where most of the languages are OV (placing the object before the verb). And in South-East Asia, and extending out into the Pacific, is a large area where VO again predominates. Other maps show much less areal patterning of this sort. Where one finds similarities of this sort within a particular geographical area, there are three sorts of explanation. One is that it is the result of contact between languages. There are a number of well-documented instances of relatively small linguistic areas, such as the Balkans and Mesoamerica, where many features are shared due to contact, irrespective of their genealogical classification. However, some of the *WALS* maps suggest the possibility of larger linguistic areas, such as one covering much of northern Eurasia.

The second possibility is that it reflects a genealogical relationship among at least some of the languages, involving a feature inherited from a common ancestor. For example, the fact that VO order predominates from Indonesia and the Philippines and extending eastward into the Pacific (but excluding much of New Guinea) reflects the fact that almost all of these languages belong to the Austronesian family and the VO order is apparently a feature that is inherited from Proto-Austronesian, the ancestor language from which all of the Austronesian languages have descended. Note that in many cases shared features within a geographical area may be partly genealogical and partly due to contact. When languages within the same family share certain features after a long period of time, contact among speakers of different languages in the family may have reinforced similarities that were features of the

protolanguage so that after a long enough period of time, the fact that the features are shared may be as much due to contact as due to their common ancestry. It should also be emphasized that features shared among languages in the same geographical area which are classified genealogically as belonging to separate families may reflect deeper genealogical connections. At first sight, this sounds like a contradiction: if they are in separate families, then how can they have a deeper genealogical connection? The answer is that it must be understood that when linguists classify languages into language families, they generally intend groups for which there is considered to be strong evidence of common ancestry. But there is little doubt that many of these families are in fact distantly related to each other but that the available evidence means that, at least at this time (and perhaps at any future time), there is a lack of convincing evidence for just which families are in fact related to which other families. But the absence of convincing evidence of a genealogical relationship does not mean that such families may not share a few features that are retentions from a common ancestor.

The third possible explanation for shared features within a particular geographical area is that it is at least partly coincidence. Especially when a map shows primarily two values, as does Map 83, with VO and OV word order, there are bound to be geographical areas in which one of these orders predominates where the occurrence throughout this area is simply accidental, where there may be two or more regions within this area which are predominantly OV or VO due to contact or for genealogical reasons, but where the fact that these regions are the same as each other is purely coincidental. Along the north coast of Papua New Guinea are a number of languages in different families, including Austronesian, which are VO. While some of these families may have acquired VO order due to contact with Austronesian, this seems rather unlikely for one family, Torricelli, since, as noted by Foley (2000: 365), most of the languages of this family are inland and have apparently had relatively little contact with Austronesian languages. Thus the contiguity of the VO languages in the Torricelli family with a few VO languages in the Austronesian family seems most likely to be coincidental.

One way to determine whether a trait that is shared within a geographical area represents nothing more than coincidence is to see whether there are other unrelated traits which are also shared among the same (or a similar) set of languages. Evidence for multiple shared traits within a particular geographical area provides evidence for a **linguistic area**. While the electronic version provides the possibility of investigating correlations among linguistic features throughout the languages of the world, it also provides the possibility of investigating correlations within particular regions. If a number of feature values correlate in a particular part of the world, but not in the world as a whole, then this is particularly strong evidence that the current distribution in that part of the world is the result of contact among its languages or possibly deep genealogical relationship. Even brief perusal of the maps will reveal, for instance, that South-East Asia often stands out from neighbouring areas, and sometimes even from the world as a whole, with the result that the materials provide a more solid basis than heretofore for the recognition of South-East Asia as a linguistic area (see, for instance, Maps 13, 51, and 55). (In addition, they show how some feature values characteristic of South-East Asia have a certain distribution beyond this area, reflecting weaker but nonetheless palpable language contact between South-East Asia and neighbouring areas with percolation of some feature values.)

Even a single trait, if relatively common in one area and relatively uncommon outside that area, can provide evidence for a linguistic area. The example cited above, of relative pronouns being common in Europe—not only in Indo-European languages but also such Uralic languages as Finnish, Estonian, and Hungarian—but uncommon elsewhere in the world, provides evidence for Europe as a linguistic area.

We hope that *WALS* will provide the interested lay reader with an overview both of the typological diversity of the world's languages —including reference to phenomena that might be judged highly unusual by the standards of the world's best-known languages—and of the ways in which that diversity is often patterned geographically, with particular areas often being characterized by particular feature values or particular combinations of feature values (whence the term "areal typology").

Although *The World Atlas of Language Structures* is a pioneering effort, we do acknowledge the efforts of earlier linguists who, on the basis of much less extensive data and with far fewer resources, nonetheless achieved significant results in mapping the world's linguistic diversity. Among our intellectual forebears, particular reference should be made to Schmidt (1926).

2 Organization

2.1 The features. *The World Atlas of Language Structures* contains 142 chapters, each consisting of a world map (plus four blow-up maps) followed by a double page of text. Each of the 142 chapters shows the distribution of a particular linguistic **feature**, reflected in the chapter's title. In a few cases, a single text accompanies a bloc of several successive maps. These texts are numbered according to the map numbers, so that a few texts receive multiple numbers (e.g. Chapters 122–123, a single text accompanying two maps).

Each chapter was contributed by an author (or team of authors) who is an expert on the particular structural feature, and who collected the worldwide cross-linguistic data from published materials and other sources. Quite a few authors are responsible for more than one chapter. Altogether 55 authors made a contribution to this work.

The 142 features are grouped thematically into the following eleven sections: phonology, morphology, nominal categories, nominal syntax, verbal categories, word order, simple clauses, complex sentences, lexicon, sign languages, and other. A complete listing of the features is provided in the Contents. As suggested by the section titles, the features span all of the major areas of language structure.

The first eight sections, phonology, morphology, nominal categories, nominal syntax, verbal categories, word order, simple clauses, and complex sentences, encompass the major structural domains of grammar. Within each of these sections, the features included provide a broad coverage of the most important subdomains within each of these structural domains. These eight sections constitute the core of the atlas, containing a total of 128 chapters.

The remaining three sections, containing a total of fourteen chapters, are of a more variegated nature. In the lexicon section, broad coverage would have been impractical; instead, a somewhat arbitrary choice of features provides a sample of the kinds of patterns that can be observed in this domain. In the sign-language section, too, the two features that are included represent arbitrary choices within the domain of sign-language morphosyntax. Finally, the last section touches on two diverse domains that are generally considered marginal to linguistic structure: paralinguistic sounds and writing systems.

The choice of features included in the atlas was governed by several competing considerations. As suggested above, the features were chosen to represent as many as possible of the major domains and subdomains of language structure. However, the desire to be as inclusive as possible came up against some practical considerations.

First, the requirement that each map represent a geographically and genealogically adequate sample of the world's languages entailed that the data for each map be obtained primarily from previously published descriptions of each individual language, typically in the form of a reference grammar. However, this limited the choice of features to those for which information is available in a typical reference grammar, and in a theory-neutral form facilitating cross-linguistic comparisons.

Secondly, the magnitude of this work, coupled with the desire to achieve a complete product within a limited time frame, entailed that the bulk of the data in the atlas derive from work already conducted over the course of a lengthy period, extending back for years and in some cases even decades. Thus, the choice of features was largely determined by what was already on offer from the contributing authors.

Due to these practical considerations, many features of current or potential future interest to linguistics had to be excluded. If the

reader is disappointed that his or her favourite feature is not to be found in the atlas, chances are that this feature simply has not been described for a sufficiently large number of diverse languages to have warranted its inclusion. Hopefully, the absence of such features from the atlas will motivate future linguists to go out and collect the necessary data.

2.2 The maps. The great majority of maps show two hundred languages or more. Map 83 ("Order of object and verb") shows the greatest number of languages (1370 languages), while the two maps on sign languages (Maps 139 and 140) show a much smaller number, for the simple reason that linguists have only recently begun to study the grammatical structure of sign languages in a comparative perspective. On average, the maps show 409 languages. This is less than 10 per cent of the world's languages,[1] so the picture that we see in this atlas is far from complete. However, not more than 10–15 per cent of languages have been described comprehensively, and many hundreds of languages are still completely or almost completely unknown. But both descriptive and comparative linguistics have made enormous progress in recent decades, and these efforts are reflected in the current work. Altogether 2559 languages, somewhat less than one half of the world's languages, occur somewhere in the atlas—we call these the ***WALS* languages**. More than 6700 books and articles have been consulted by the authors and the relevant bibliographical references can be accessed in the electronic version of the atlas. In addition to the maps and accompanying texts, the atlas contains a genealogically organized list of the 2559 languages (the Genealogical Language List, by Matthew S. Dryer, on p. 584), with additional information on geographical location and alternative names, to facilitate identification of each language. At the end of the atlas is a language (and language family) name index. Issues having to do with the identification and designation of languages and language families are discussed in detail in §3 below.

The two map pages for each feature consist of a single world map (omitting unpopulated Antarctica) and below it four blow-up maps of areas with a particularly high density of languages. The world map is Pacific-centred, a representation that will look unfamiliar to many readers from Europe, Africa, and the Americas. A Pacific-centred map was chosen for two reasons. First, unlike the Atlantic Ocean, the Pacific Ocean has many small islands on which different languages are spoken. A Europe–Africa-centred map would have meant showing some Polynesian languages on the right and others on the left. Second, human populations spread out of Africa into Eurasia and across the Bering Strait into the Americas. Some linguists have suggested that similarities between Eurasian and American languages to some extent still reflect ancient population movements, and these similarities can only be captured in a Pacific-centred map. Contacts across the Atlantic have been much too recent to have an effect on the linguistic picture as the atlas shows it.

The world map is in Robinson projection, a projection which distorts shape, area, scale, and distance to create attractive average projection properties. The scale is 696 km per cm at the equator, 606 km per cm at 30 degrees south or north of the equator, and 352 km per cm at 60 degrees south or north of the equator. The shapes of regions north and south of 45 degrees latitude (including much of Europe and North America) are significantly distorted in this projection. However, for the purposes of displaying the world's languages, traditional rectangular Mercator maps would be unsuitable because they compress the areas around the equator, which are precisely the areas with the highest density of languages (see Nettle 1999 for documentation and discussion of this fact). Moreover, the ellipse in the Robinson projection better reflects the essentially round character of the earth than rectangular maps.

The four blow-up maps below the world map show exactly the same languages and symbols as the main map, but because of the greater resolution of the maps, many language dots that are covered by other dots on the main map are now distinguishable. Many chapters provide information on so many languages that even on the blow-up maps, many language dots are still underneath other language dots and hence invisible on the printed maps. Atlas users can turn to the interactive electronic version and zoom in closely enough that all languages can be identified. Which language dots are on top of others and thus visible in the printed version is largely random, though we have tried to make sure in some cases that languages which are of particular interest for that map (e.g. those that are mentioned in the accompanying text) will not be hidden underneath other languages.

Three chapters are somewhat special with respect to the maps. The two chapters on sign languages (Chapters 139–140) show pictures of signs instead of blow-up maps, and the chapter on writing systems (Chapter 141) shows differently coloured areas rather than differently coloured dots, as well as specimens of written texts in different systems instead of blow-up maps.

2.3 The feature values. Each feature is associated with a set of **feature values**, forming the basis for distinguishing between languages of different types. The simplest maps show just two different values. For example, Map 107 on passive constructions shows two language types, those possessing a passive construction and those lacking a passive construction. Most maps distinguish between three and five values, but some distinguish up to nine values (Chapters 33, 49, and 51). Although language types can usually be subdivided into many more subtypes, the values have been limited to at most nine in *WALS* because more types are difficult to represent by different colours, and users would find them hard to distinguish anyway. Each value is represented by a unique symbol, most often characterized by a particular colour, but sometimes also by a particular shape. Each symbol contains a three-letter code (the ***WALS* code**) that uniquely identifies the language (see §3.5 below, and the *WALS* Code Index at the end of the atlas).

Within each chapter, the feature values are indicated in abbreviated form in a small box on the map itself (in the southern Indian Ocean), and in expanded form in a larger box (called the **feature-value box**) within the accompanying text. For each feature value, these two boxes also show the colour and shape of the associated dot, and the number of languages characterized by the feature value in question. In addition, the text of each chapter provides a detailed description of the set of feature values, including the criteria that were used to assign feature values to individual languages.

One general requirement on the feature values of each chapter is that they be exhaustive. What this means is that for each feature, each and every language under consideration must be assigned one of the feature values; there are no "blank" cases of languages that do not have a feature value. In order to meet this requirement, many chapters include one or more feature values that account for various cases involving nonapplicability, indeterminacy, or mixing. For example, many maps characterize the different kinds of a certain construction, such as the associative plural (Chapter 36), but also include an additional feature value corresponding to the absence of the construction in question. Similarly, many of the word order maps (Chapters 81–97), such as noun–adjective order, also include an additional feature value corresponding to the absence of a dominant word order.[2]

Each feature value is associated with a dot of distinctive colour and, in some maps, also shape. The choice of symbols is intended to reflect, as accurately as possible, the logic underlying each set of feature values. For example, if feature values form an ordered set, then their symbols should ideally also form an ordered set. Similarly, feature values that are conceptually closer to each other should ideally be associated with symbols that are closer to each other. In many cases, however, it proved impossible to come up with a set of symbols that would be completely isomorphic with the logic of the feature values. Nevertheless, the reader will notice a number of colour and shape schemes which recur throughout the atlas,

[1] According to *Ethnologue* (Grimes 2000, vol. 1, p. 846), there are currently 6809 living languages.

[2] However, the exhaustiveness requirement is violated in Maps 14–17 on stress, which are silent on languages lacking stress. There was no way around this, because these chapters are based on a database collected earlier (StressTyp) that did not include stressless languages.

underscoring the presence of particular logical patterns each of which is shared by a variety of different maps.

No attempt has been made to make chapters by different authors that overlap in their features consistent, and it is not hard to find inconsistencies between certain chapters. For example, there are chapters by three different authors (or sets of authors) which deal with nominal case. There are some languages which are shown as lacking case by one author and as having case by a second author. These inconsistencies can arise for a number of reasons. Some occur because different authors use different criteria for identifying case, these differences hopefully being clear from the text accompanying the maps. Some occur because different authors classify particular borderline cases differently. Some occur because different authors use different sources and the sources describe the phenomenon differently. And some occur simply because of errors, either in an author's interpretation of their source or simply due to a coding error. In fact, quite independently of inconsistencies between chapters, there are bound to be errors of coding that will only be discovered after this atlas has been published. The editors will keep a **catalogue of reported errors** on the internet at <http://www.wals.info/>

3 The languages

3.1 The WALS samples. There is a total of 2560 languages which appear on at least one map in the atlas. Some of these languages (265 in number) appear on only one map, while some, such as English, appear on most of the maps. There are 175 languages which appear on at least eighty maps, and 424 languages which appear on at least forty maps. The choice of which languages to include on particular maps was the choice of individual authors. However, there is a set of 100 languages (hereafter the 100-language sample) which authors were asked to include on their maps if at all possible, and a further 100 languages which authors were encouraged to include on their maps (hereafter these two sets of 100 languages together are referred to as the 200-language sample).

A general desideratum for a good language sample is that it maximize both genealogical and areal diversity. Samples which include too many languages from one area of the world or too many languages from one family can provide a misleading picture of the relative frequency of different types of languages. Typological studies in the past have often included a disproportionate number of Indo-European languages or of languages of Europe or Eurasia. While Eurasia has a larger landmass than any other continental region in the world, fewer than 20 per cent of the languages of the world are spoken on mainland Eurasia (i.e. excluding the languages of Indonesia and the Philippines and other islands). In fact, there are more languages spoken on the island of New Guinea than in mainland Eurasia. Furthermore, as a number of the maps in this atlas show, there are patterns of similarity among languages of Eurasia that one does not find elsewhere in the world. For example, Map 97 shows that the vast majority of the OV languages of Eurasia (i.e. ones that place the object before the verb) place the modifying adjective before the noun. From this, linguists in the past erroneously concluded that this was a normal feature of OV languages. But as Map 97 shows, this is not true outside of Eurasia, where OV languages more often place adjectives after the noun. Only by using samples of languages which include many languages from outside Eurasia can we avoid making erroneous inferences of this sort.

Maximizing genealogical and areal diversity were major considerations in constructing the 100- and 200-language samples. However, there were a number of other considerations that played a role in constructing these samples that would not generally play the same role in constructing samples of languages. First, most of the languages of the islands of the Pacific fall within the Oceanic branch of the Austronesian family and thus are closely related to each other. For instance, one would normally not include more than one of these languages in a sample of 100 or even 200 languages. However, because the sample used here is for an atlas, we decided that we ought to include more of these languages, since otherwise there would be few dots on the maps in the Pacific. For this reason, there are two Oceanic languages in the 100-language sample and seven in the 200-language sample. Similar considerations led to the inclusion of three Bantu languages in the 100-language sample and five in the 200-language sample. Without these, many of the maps would have shown few languages in sub-Saharan Africa, and the majority of those shown would have been non-Bantu languages that are in some ways atypical of this region. A second consideration that would not normally play a role in constructing a language sample is that we felt that we ought to include a number of the major languages of Eurasia, even when this meant including pairs of languages which are too close genealogically to be otherwise included in a sample of 100 or 200 languages, including English and German, French and Spanish, and Modern Hebrew and Egyptian Arabic.

A further consideration in choosing languages for the 100- and 200-language samples was the ready availability of detailed grammatical descriptions. In most cases, the choice of a language over genealogically related languages was based on the availability of detailed descriptions. Some of the languages that were included in the samples are ones for which there is no detailed description but for which an expert on the language was willing to answer questions from authors (see §4 below). Some of the languages in the 200-language sample were chosen primarily for the purposes of maximizing genealogical or areal diversity, despite the fact that the available descriptions of these languages are somewhat meagre, thus making it impossible for many authors to include them on their maps. One language in the 200-language sample, Minica Huitoto, appears on only 32 maps; however, this was because we eventually realized the need to distinguish this language from other Huitoto languages and some authors in attempting to include Huitoto used sources for one of these other languages.

The choice of which languages to include in the 100-language sample and which to include in the 200-language sample was based on the following considerations. Languages which are more well known were normally placed in the 100-language sample. Languages with more readily available or more detailed descriptions were also placed in the 100-language sample. However, we also attempted to maximize genealogical and areal diversity in both samples so that in some instances, a language was placed in the 200-language sample but not in the 100-language sample if there was already a language that was close genealogically or geographically in the 100-language sample.

The following is a list of the languages in the 200-language sample, with those in the 100-language sample marked with an asterisk. The languages are organized in the same fashion as the entire set of *WALS* languages in the Genealogical Language List, by family, subfamily and genus (see the introduction to the Genealogical Language List on p. 584 for an explanation of the notion of genus). Names of genera are placed in italics.

Khoisan:
- *Central Khoisan*: ***Khoekhoe**
- *Northern Khoisan*: **Ju|'hoan**

Niger-Congo
- *Adamawa-Ubangian*: ***Sango**
- Atlantic
  - *Northern Atlantic*: **Diola-Fogny**
- Benue-Congo
  - *Bantoid*: **Kongo, *Luvale, Nkore-Kiga, *Swahili, *Zulu**
  - *Defoid*: ***Yoruba**
  - *Igboid*: **Igbo**
- *Gur*: ***Supyire, Koromfe**
- *Kru*: ***Grebo**
- *Kwa*: **Ewe**
- Mande
  - *North-West Mande*: **Bambara**

Kadugli: ***Krongo**

Nilo-Saharan
- Central Sudanic
  - *Bongo-Bagirmi*: ***Bagirmi**
  - *Lendu*: **Ngiti**

Eastern Sudanic
Nilotic: ***Lango**
Nubian: **Dongolese Nubian**
Surmic: **Murle**
Fur: **Fur**
Kunama: **Kunama**
Maban: **Maba**
Saharan: **Kanuri**
Songhay: ***Koyraboro Senni**

Afro-Asiatic
Berber: ***Middle Atlas Berber**
Chadic
East Chadic: **Kera**
West Chadic: ***Hausa**
Cushitic
Beja: **Beja**
Eastern Cushitic: ***Harar Oromo**
Southern Cushitic: **Iraqw**
Semitic: ***Egyptian Arabic, *Modern Hebrew**

Basque: ***Basque**

Indo-European
Armenian: **Eastern Armenian**
Baltic: **Latvian**
Celtic: **Irish**
Germanic: ***English, *German**
Greek: ***Modern Greek**
Indic: ***Hindi**
Iranian: ***Persian**
Italic: ***French, *Spanish**
Slavic: ***Russian**

Uralic
Finno-Ugric
Finnic: ***Finnish**
Ugric: **Hungarian**
Samoyedic: **Nenets**

Altaic
Mongolic: ***Khalkha**
Tungusic: **Evenki**
Turkic: ***Turkish**

Yukaghir: **Kolyma Yukaghir**

Yeniseian: **Ket**

Chukotko-Kamchatkan: ***Chukchi**

Nivkh: **Nivkh**

Ainu: **Ainu**

Japanese: ***Japanese**

Korean: ***Korean**

North-West Caucasian: ***Abkhaz**

Nakh-Daghestanian
Daghestanian
Avar-Andic-Tsezic: **Hunzib**
Lak-Dargwa: **Lak**
Lezgic: ***Lezgian**
Nakh: **Ingush**

Kartvelian: ***Georgian**

Burushaski: ***Burushaski**

Dravidian
Dravidian Proper: ***Kannada**
North-West Dravidian: **Brahui**

Sino-Tibetan
Chinese: ***Mandarin**
Tibeto-Burman
Baric: **Garo**
Bodic: **Ladakhi**
Burmese-Lolo: ***Burmese**
Karen: **Kayah Li**
Kuki-Chin-Naga: **Bawm, *Meithei**
Lepcha: **Lepcha**

Hmong-Mien: ***Hmong Njua**

Tai-Kadai
Kam-Tai: ***Thai**

Austro-Asiatic
Munda: **Mundari**
Mon-Khmer
Aslian: **Semelai**
Khasi: **Khasi**
Khmer: **Khmer**
Palaung-Khmuic: **Khmu'**
Viet-Muong: ***Vietnamese**

Austronesian
Paiwanic: ***Paiwan**
Eastern Malayo-Polynesian
Oceanic: **Drehu, *Fijian, Kilivila, Kiribatese, Maori, Paamese, *Rapanui**
South Halmahera-North-West New Guinea: **Taba**
Western Malayo-Polynesian
Borneo: ***Malagasy**
Chamorro: ***Chamorro**
Meso-Philippine: ***Tagalog**
Sulawesi: ***Tukang Besi**
Sundic: **Karo Batak, *Indonesian**

West Papuan:
North-Central Bird's Head: ***Maybrat**

Sentani: **Sentani**

Border: ***Imonda**

Torricelli
Kombio-Arapesh: ***Arapesh**

Sepik
Sepik Hill: ***Alamblak**

Lower Sepik-Ramu
Lower Sepik: **Yimas**

Trans-New Guinea
Angan: **Hamtai**
Asmat-Kamoro: ***Asmat**
Binanderean: **Suena**
Dani: ***Lower Grand Valley Dani**
Engan: ***Kewa**
Madang: ***Amele, Kobon, Usan**
Mek: **Una**
Wissel Lakes-Kemandoga: **Ekari**

Marind
Marind Proper: **Marind**

Dagan: ***Daga**

Solomons East Papuan: ***Lavukaleve**

Australian
Bunuban: ***Gooniyandi**

Daly
Western Daly: **Maranungku**

Gunwinyguan
Nunggubuyu: **Nunggubuyu**
Yangmanic: **Wardaman**
Iwaidjan: ***Maung**
Mangarrayi: ***Mangarrayi**
Pama-Nyungan: ***Martuthunira, *Ngiyambaa, Pitjantjatjara, Yidiny**

Tangkic: ***Kayardild**
Tiwian: ***Tiwi**
West Barkly: **Wambaya**
Wororan: **Ungarinjin**

Eskimo-Aleut: ***West Greenlandic, Central Yup'ik**

Na-Dene
Athapaskan: **Navajo, *Slave**
Tlingit: **Tlingit**

Haida: **Haida**

Algic
Algonquian: ***Plains Cree, Passamaquoddy-Maliseet**
Yurok: **Yurok**

Iroquoian: ***Oneida**

Yuchi: **Yuchi**

Muskogean: ***Koasati**

Tunica: **Tunica**

Caddoan: ***Wichita**

Siouan: ***Lakhota**

Kiowa-Tanoan: ***Kiowa**

Keresan: ***Acoma**

Uto-Aztecan
Aztecan: **Tetelcingo Nahuatl**
Cahita: ***Yaqui**
Numic: **Comanche**
Takic: **Cahuilla**

Wakashan: ***Makah**

Salishan: **Squamish**

Kutenai: ***Kutenai**

Sahaptian: **Nez Perce**

Penutian
Miwok: **Southern Sierra Miwok**
Tsimshianic: **Coast Tsimshian**

Oregon Coast
Coosan: **Hanis Coos**

Hokan
Pomoan: **Southeastern Pomo**
Yuman: ***Maricopa**

Karok: ***Karok**

Oto-Manguean
Chinantecan: **Lealao Chinantec**
Mixtecan: ***Chalcatongo Mixtec**
Otomian: ***Mezquital Otomí**

Mixe-Zoque: ***Copainalá Zoque**

Mayan: ***Jakaltek**

Chibchan
Aruak: **Ika**
Rama: ***Rama**
Talamanca: **Bribri**

Choco: **Epena Pedee**

Barbacoan: **Awa Pit**

Tucanoan: ***Barasano**

Huitotoan: **Minica Huitoto**

Warao: ***Warao**

Yanomam: ***Sanuma**

Peba-Yaguan: ***Yagua**

Panoan: **Shipibo-Konibo**

Quechuan: ***Imbabura Quechua**

Aymaran: **Aymara**

Arawakan: ***Apurinã**

Cariban: **Carib, *Hixkaryana**

Tupian
Tupi-Guaraní: ***Guaraní, Urubu-Kaapor**

Macro-Ge
Ge-Kaingang: ***Canela-Krahô**

Trumai: **Trumai**

Chapacura-Wanhan: ***Wari'**

Mura: ***Pirahã**

Arauan: **Paumarí**

Tacanan: **Araona**

Cayuvava: **Cayuvava**

Matacoan: ***Wichí**

Guaicuruan: **Abipón**

Araucanian: ***Mapudungun**

Alacalufan: **Qawasqar**

Chon
Chon Proper: **Selknam**

Creoles and Pidgins: **Ndyuka**

While the 100- and 200-language samples could be used as samples for other typological studies, a genealogically more balanced sample, with only one language per genus, would remove Kongo, Luvale, Nkore-Kiga, Zulu, Koromfe, Modern Hebrew or Egyptian Arabic, German or English, French or Spanish, Bawm, Drehu, Kilivila, Kiribatese, Maori, Paamese, Rapanui, Karo Batak, Kobon, Usan, Ngiyambaa, Pitjantjatjara, Yidiny, Central Yup'ik, Navajo, Passamaquoddy-Maliseet, Carib, and Urubu-Kaapor. Three languages (Maybrat, Makah, Kutenai) were included because of the availability of language experts to answer questions from chapter authors or because copies of unpublished descriptions were made available to authors, and therefore might not be included in a sample used for future typological studies.

3.2 Language versus dialect. There are a number of instances in which distinct dialects are distinguished in the atlas. For example, *WALS* distinguishes four dialects of Inuktitut, the Eskimo-Aleut language of northern Canada. In a more extreme case, Map 54 shows a number of different dialects of German (because dialects vary with respect to the phenomenon investigated). In other cases, what is shown as a language is probably in fact a set of closely related but mutually unintelligible languages. An example of this is Bikol, spoken in the Philippines. No systematic attempt has been made here to distinguish dialects of the same language from different languages. Where authors submitted separate data for more than one dialect of a language, we generally maintained the distinction, and where authors submitted data using a name that covers a number of mutually unintelligible languages, we endeavoured to ascertain which language their data was based on, but in some cases we were unable to do so, and in some cases, their data was based on more than one source, where the sources describe different varieties.

We originally attempted to have a set of *WALS* languages in which no language was a variety of another *WALS* language. While we were able to minimize this, it proved impossible to apply the principle consistently. For example, as just mentioned, one map shows a number of varieties of German, while all other maps that include German simply show *German*, without specification of a particular variety. A more typical example is provided by Irish. Here, most authors provide data for Irish, without specifying a variety. But one author submitted data for Donegal Irish and another author data for Munster Irish (since the dialects differ from each

other in some ways). The *WALS* set of languages also occasionally includes pairs in which one is a variety of the other, where the relationship is one of language to set of closely related languages rather than one of dialect to language. For example, while we generally distinguish different Huitoto languages, a few authors submitted data where it was not clear which Huitoto language their data was based on and for their maps, the *WALS* language is simply called *Huitoto.*

3.3 Language names. Many languages are known in the literature under different names. We have attempted to choose names for *WALS* that are the names by which the languages are currently known. This means that our name is occasionally different from that used in some sources on the language, where we have reason to believe that the name used in the source does not conform to more recent usage. For example, older sources on O'odham refer to the language as *Papago*, but we use the more recent name *O'odham*. Older names are often considered offensive by communities in which the language is spoken. In the Genealogical Language List on p. 584 we also give the name of the language as it is listed in *Ethnologue*, as best we can determine. The Language Name Index also includes the *Ethnologue* names as well as other names used in sources for the language (such as *Papago*).

When names of languages involve two or more words with a modifier followed by a head in the usual English name for the language, where the head denotes a language or language group and the modifier identifies a particular variety of that language or language group, we have two ways of referring to the language. In the chapter texts we use the usual name, with the modifier preceding the head, but in the Genealogical Language List and in the electronic version, we place the head first with the modifier following in parentheses. Thus, what is called *Chalcatongo Mixtec* in chapter texts is called *Mixtec (Chalcatongo)* in the list and in the electronic version. Note that we follow this convention both when the head denotes a language and the modifier identifies a dialect of that language (e.g. *Irish (Donegal)*) and when the head denotes a language group and the modifier identifies a particular language in that group (e.g. *Mixtec (Chalcatongo)*).

There are a number of cases that may look like instances of this, but where the head does not denote a language or language group of which the modifier identifies a variety. For example, Upper Kuskokwim remains in that form because it is not the upper variety of a Kuskokwim language (there is no Kuskokwim language or language group); rather, it is a language spoken in the vicinity of the Upper Kuskokwim river. Similarly, Tümpisa Shoshone remains in that form since it is not a variety of Shoshone, but just a closely related language. Some language names in English already occur in the form Head Modifier because their name reflects the syntax of some other language in which the modifier follows the head. These remain in the form Head Modifier, without parentheses. An example of this is Hmong Daw. Also we retain the order Modifier Head if the Head denotes a type of language, such as creole, pidgin, or sign language, rather than a genealogical category. An example of this is Berbice Dutch Creole.

There are many instances of homophonous language names, where two languages in different parts of the world happen to have the same name. Where two such languages occur in the set of *WALS* languages, we generally disambiguate them by adding a modifier of the form *in* plus country name in parentheses, as in *Baka (in Cameroon)* and *Baka (in Sudan)*, where the former is a Niger-Congo language and the latter is a Nilo-Saharan language. We sometimes employ this usage for only one language, where we use a different name for the second language, simply because the ambiguous name is sometimes used by others as a name for the second language. An example of this is *Mono (in United States)*, where *Mono* is also the name used by *Ethnologue* for the Austronesian language we call *Mono-Alu*. Occasionally this approach does not suffice to disambiguate a language name when both languages are spoken in the same country. In this case, we place the name of the language family in parentheses. An example of this is *Motilón (Chibchan)*, where the family is needed to distinguish it from the Carib language Yukpa, which is sometimes known as *Motilón* and which is also spoken in Colombia and Venezuela. Note that there are many instances of homophonous language names where we do not add a modifier because the other languages with this name are not *WALS* languages. For example, *Ethnologue* lists two languages called *Bulu*, one spoken in Cameroon, the other in Papua New Guinea. Since only the former is a *WALS* language, we simply call this language Bulu.

While our names include standard diacritic symbols (as in *á*, *ä*, or *ã*), we avoid using superscripts or symbols that are not standard symbols. Thus we use *Yidiny*, rather than *Yidinʸ* or *Yidiɲ*.

3.4 Locations of languages. The languages are represented on the maps as dots, rather than as regions, but it should be borne in mind that many languages are spoken over areas larger than the dots. We attempted to locate the dots somewhere near the centre of the region where the languages are spoken, although in some cases this was difficult because the region in which the language is spoken is discontinuous. In these cases, we generally located the dot within the larger region in which the language is spoken. In some cases, the location of the dot is based on the location of a major city, town, or village in which the language is spoken. For example, Egyptian Arabic is located in the vicinity of Cairo, rather than in the middle of Egypt. For most languages, the location of the dots is based on their location on maps in Moseley and Asher (1994) or *Ethnologue* (Grimes 2000). For languages spoken in Canada and the United States, the location is based on Goddard (1996). For languages spoken in Nepal, the location is based on maps in Bradley (1997). For many of the languages spoken in Australia, the location is based on Tindale (1974). In some instances, the location is simply based on an explanation in the specific sources for the language. Future work will probably make clear that the location of some of the dots is inaccurate. Where we find inaccuracies after *WALS* is published, we will post information at http://www.wals.info/.

Note that in identifying the location of languages, we use locations prior to European colonial expansion. This means that the dot for English is located in England, and not in some other country where English is spoken. Similarly, Spanish is located in Spain, despite the fact that the majority of speakers are in the Americas. Thus, the languages shown in the Americas are for indigenous languages and for creoles and sign languages (since instances of the latter two types in the Americas only have locations in the Americas). Analogously, indigenous languages that are now spoken in locations different from where they were spoken at the time of European contact are located in their location at the time of European contact rather than their present location (in contrast to the practice in *Ethnologue*). For example, a number of indigenous languages of the United States are now only or primarily spoken on reservations in Oklahoma, often far from where their speakers originally lived. An example of this is Yuchi, originally spoken much further east, in what is now Tennessee.

3.5 Three-letter codes. The languages are identified on the maps by means of a three-letter code (***WALS* code**). A list of the *WALS* codes and the languages they denote is given at the end of this atlas. Readers who find these three-letter codes difficult to read will find them easier to identify on the zoomable maps of the elctronic version. We considered using the three-letter codes employed by *Ethnologue* but decided not to, both because there are many instances in which the languages in the atlas either represent varieties of languages in *Ethnologue* or correspond to an entire set of languages in *Ethnologue*, and because we wanted to use three-letter codes that are more mnemonic. The Genealogical Language List on p. 584 does give the *Ethnologue* three-letter code for each *WALS* language, as best we can determine. In assigning three-letter codes to languages, we first attempted to use the first three letters of the language name, unless the language name includes two or more words, in which case we attempted to use three letters based on initial letters in the different words. In other cases (where there is more than one language with the same three initial letters), we used the first three consonants in the name of the language (ignoring vowels). In some cases we had to use some other sequence of letters appearing somewhere in the name of the language and in a few instances we had to resort to

adding as the third letter some letter that does not appear in the name of the language at all. But we always use the first letter in the *WALS* language name as the first letter in the three-letter code. With language names which are represented as *Modifier Head* in the chapter texts but as *Head (Modifier)* in the Genealogical Language List and in the electronic version, we use the latter for determining the three-letter code so that the first letter in the three-letter code matches the first letter in the head. For example, the three-letter code for Chalcatongo Mixtec, called *Mixtec (Chalcatongo)* in the Genealogical Language List and in the electronic version, is *mxc*. For languages whose names consist of two letters, such as Ik, we use a two-letter code rather than a three-letter code.

4 The data sources

The maps of the *World Atlas of Language Structures* are largely based on published primary sources that provide information about the languages in question. These include full grammars and dictionaries, but also more specialized articles that are confined to particular aspects of the language structure (e.g. only the phonological structure, or only certain syntactic constructions). Unpublished dissertations have also been used as sources, because these are often accessible to typologists. In a few cases (for languages like English, Spanish, or Russian), the authors have relied on their own knowledge of the language. Secondary sources, i.e. published typological surveys based on primary sources, have also occasionally been used where it seemed hard to avoid, although it is now generally recognized that comparative linguists should ideally work with primary sources.

Many dialect atlases have worked with an entirely different method of data collection, based on questionnaires. These are drawn up by the atlas editors and filled in by different fieldworkers on location for each data point. This is practical for dialect atlases because the fieldwork is restricted to a relatively small area, but for an atlas of global scope like *WALS*, this method would have required a budget a thousand times larger. Moreover, the questionnaire method is problematic for more sophisticated cross-linguistic work, because identifying certain phenomena (e.g. iambic rhythm or applicative constructions) in different languages requires detailed knowledge of the phenomenon in question.

The editors briefly considered the possibility of basing the atlas on questionnaires sent to a set of experts who know their respective languages so well that they would be able to answer structural questions without additional fieldwork. However, it quickly became clear that while this method would have the advantage of showing a uniform sample of languages on all maps (as in dialect atlases), it would be difficult to find enough experts willing to collaborate on such an enterprise, and it would not be wise to leave the rich data sources of published descriptions untapped.

As was mentioned in §3.1, the editors encouraged the authors to try and provide data on a fixed sample of 100 core languages and a further sample of 100 additional languages. For quite a few of these languages, the editors contacted experts, asking them to serve as consultants for the *WALS* authors; a list of their names is given in the Acknowledgements. Some authors also contacted other experts and received relevant data by personal communication from them.

More than 58,000 data points are shown on the *WALS* maps. Of these, about 2000 (or 3.4 per cent) are based on personal communications from experts, about 700 (or 1.2 per cent) are based on the authors' own knowledge or own data, and about 400 (or 0.7 per cent) are based on secondary sources.

For each data point (i.e. language-feature pair), the source used by the author is given in the interactive electronic version. For many of the data points, the electronic version also provides an example of the phenomenon in question.

5 The Interactive Reference Tool

The interactive electronic version (on CD-ROM) of the *World Atlas of Language Structures* contains the entire database on which this atlas is based and allows the user to display the data in a variety of ways, to conduct automatic searches, to export data and maps, and to create compound features based on the standard 141 features of the printed version.

The data in the electronic version of the database should be thought of as an appendix to *WALS* and its individual chapters. Thus, scholars who make use of these data in their research must refer to all the relevant chapters and give credit to their authors. It is not sufficient just to refer to "the electronic version of *WALS*".

Users of the electronic version can customize the map in various ways: show major cities and country names, remove country boundaries and rivers, and replace the light green/light blue base map by a topographic map showing altitude levels. The language dots can be shown in five different sizes, and the language name can be shown either as the three-letter *WALS* code inside the symbol, or in full to the right of the symbol. The colours and shapes of the symbols can be changed. When the mouse pointer moves over the dot, the full name is shown, and when clicking on a dot, a window with further information on the language opens (including the data source). Users can also zoom in on areas with high dot density, closely enough to see all dots separately, and drag on a map to see adjacent areas. Maps can be exported and printed, and various user-defined selections can be saved for future use. In some chapters, an example is provided for each data point.

The electronic version allows users to manipulate the standard features in two ways: values can be removed (if they are not of interest in a certain context), and several values can be merged into a single value. For instance, the five values of Chapter 1 (small, moderately small, average, moderately large, large) can be reduced to three (below average, average, above average) with just two mouse clicks.

Users can search for language names, genus names, family names, country names, and even for text in bibliographic entries. It is possible, for instance, to find and display all languages beginning with *X*, all languages belonging to the Austronesian family, all languages spoken in Colombia, or all languages described by Jeffrey Heath. On the electronic maps that only show languages (without giving information about the features), different dot colours may stand for different families or different genera.

Most importantly for comparative linguists, users can create their own compound features. For example, a linguist may want to know whether the existence of tone in a language is correlated with the type of syllable structure. Both features have three values (tone: none, simple, complex; syllable structure: simple, moderately complex, complex), so by combining them, one gets nine possible values:

No tones AND Simple syllable structure
No tones AND Moderately complex syllable structure
No tones AND Complex syllable structure
Simple tone system AND Simple syllable structure
Simple tone system AND Moderately complex syllable structure
Simple tone system AND Complex syllable structure
Complex tone system AND Simple syllable structure
Complex tone system AND Moderately complex syllable structure
Complex tone system AND Complex syllable structure

The program automatically creates a compound feature with these nine values, shows the number of languages for each value, suggests a symbol for each value, and displays a map of the compound feature. More complex ways of creating compound features are also possible and are described in detail in the electronic version.

6 Disclaimer

In identifying the status of speech varieties as languages or dialects, in assigning names to languages and dialects, in identifying countries, and in locating languages in countries, we have been guided solely by practical considerations and by current scholarly practice. In no instance should our usage be taken as implying a particular political stance or as insulting the speakers of a particular speech variety.

TEXTS AND MAPS

1. Consonant Inventories

IAN MADDIESON

1 Introduction

This chapter and the next few chapters will look at various aspects of the complexity of the sound resources used in the world's languages and examine how this complexity is distributed geographically. The first aspect to be examined is the size of the set of consonants used in the language, usually referred to as the **consonant inventory**. This is one element of what is called the **phonology** of the language.

1.	Small	91
2.	Moderately small	121
3.	Average	181
4.	Moderately large	116
5.	Large	53
	total	562

It is usually possible to agree for any given language on a set of elements which are considered to be the **speech sounds** used in that language. The most important consideration in deciding on this set is to find groups of words which sound different from each other by the smallest degree sufficient to make them distinct words of the language. For example, the English one-syllable words *pin*, *tin*, *kin*, *fin*, *thin*, *sin*, *shin* are part of a set which differ by beginning in different ways, *dim*, *din*, *ding*, *did*, *dig*, *dish* are part of a set which differ by ending in different ways and *pin*, *pen*, *pan*, *pun*, *pain*, *pine*, *pawn* are part of a set which differ in the middle of the syllable. From a series of such comparisons a list of candidate speech sounds for the language will emerge. Generally the set of those which can appear at the beginnings and ends of syllables will be unlike those which can occur in the middle, hence a distinction is made between **consonants** (sounds typically occurring at the syllable margins) and **vowels** (sounds typically occurring in the syllable centres). In this chapter only consonants will be discussed.

Several further decisions must be made, such as which consonants in different positions should be considered to be the same as each other. For example, speakers of English generally consider that words such as *pip*, *tit*, *kick*, *bib*, *did*, *gig* begin and end with the same consonant even though there are some easily recognizable differences between the sounds at the beginning and those at the end. It is also necessary to resolve questions about whether certain beginnings or endings of syllables should be considered to be one sound or a sequence of two or more sounds when analysed from the point of view of the structure of the particular language. For example, the English word *chip* begins in a way that is similar to the beginning of *tip* followed by the beginning of *ship* (compare saying *grey chip* and *great ship*), and the English word *quick* begins in a way that is similar to the beginning of *kick* followed by the beginning of *wick* (compare saying *lie quick* and *like wick*). These syllable beginnings would both be noted in a phonetic transcription with two symbols, as /ʧ/ and /kw/ respectively. However, when we consider the possibilities of finding related sequences in English, a difference between the two becomes apparent. Nothing except /t/ can precede /ʃ/ at the beginning of an English syllable, whereas other sounds can precede /w/, as in *twin*, *swim*, *dwell*, *thwart*. Also several other sounds can follow /k/, as in *click*, *crick*, suggesting that /k/ and /w/ in the /kw/ sequence are independent elements. Although words like *trip*, *twin* might suggest independence of the parts /t/ and /ʃ/ in *chip*, the sequences /tw, tr/ are not similar to /ʧ/ in an important way. This is because no English syllable can end with /tw, tr/ (or with /kw, kl, kr/), whereas syllables can end with /ʧ/, as in *rich*, *pitch*, *kitsch*. These considerations suggest that /ʧ/ is behaving like a single consonant in English, whereas /kw/ is a sequence of two separate consonants.

When such decisions have been made, a list of the consonants used in the language can be compiled and the total of distinct ones added up. For English, there is general agreement that the consonant inventory contains 24 consonants, though some linguists might decide there are one or two more or less than this. In the survey of 562 languages reported here a strong effort has been made to apply consistent criteria in determining the consonant inventory size. This sometimes leads to some difference from the conclusions in published descriptions of the languages concerned. For most languages relatively straightforward decisions can be reached, but others are more problematic. A difficult choice often concerns whether to include consonants found only in words borrowed from other languages; generally those sounds introduced just in the last few generations as the result of the spread of world languages such as English, Spanish, Russian, Mandarin, and Modern Standard Arabic have been excluded.

The range of resulting inventories extends from a low of six consonants to a high of 122. Rotokas (West Bougainville; Papua New Guinea) has only six consonants. These might be represented in a simplified transcription with the letters /p, t, k, b, d, g/ although the range of pronunciations heard in different word positions covers a considerably wider range of sounds than these letters suggest. !Xóõ (Southern Khoisan; Botswana) has 122 consonants, mainly because it has a very large number of different click sounds with which a word may begin. The more typical consonant inventory size is in the low twenties, with the mean for the 562 languages being 22.7, the modal value 22, and the median 21. Consonant inventories close to this size (22 ± 3) have been categorized as **average**, and the remainder divided into the categories **small** (from 6 to 14 consonants), **moderately small** (15–18), **moderately large** (26–33), and **large** (34 or more consonants). As Figure 1 illustrates, the particular cut-off values for the categories were chosen so as to approximate a histogram with a normal distribution, although there are somewhat more languages with inventories smaller than the band defined as "average" than with larger than average inventories.

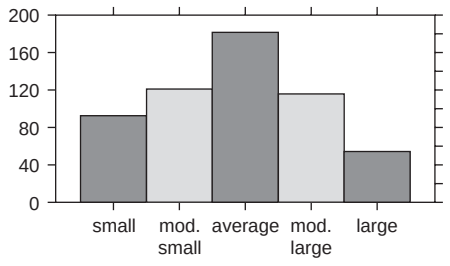


Figure 1 Histogram of languages in the sample according to categories of consonant inventory size

2 Geographical distribution

Languages with average-size consonant inventories are found in most areas of the world, suggesting that this size truly is representative of something typical for spoken human languages. The languages with larger or smaller inventories on the other hand display quite marked regional disparities in their distribution.

Those with smaller than average consonant inventories predominate in the Pacific region (including New Guinea), in South America, and in the eastern part of North America, with particular concentrations of "small" inventories in New Guinea and the Amazon basin. The degree of typological similarity with respect to consonant inventory size between the languages of New Guinea and Australia is intriguing. The received idea is that the population

ancestral to speakers of today's Australian languages reached the continent when New Guinea and Australia were connected by dry land in the now partly submerged landmass known to geologists as the Sahul shelf. Since the land bridge linking New Guinea and Australia was severed around 7000 years ago, contact between Australian and New Guinea peoples is believed to have been strictly limited except in the immediate region of the Torres Straits. Could this similarity represent the conservation of a trait common to languages spoken long ago when the lands were joined?

Those with larger than average consonant inventories are strongly represented in Africa, especially south of the equator and east of the Bight of Benin; smaller inventories occur generally in the west. Languages with larger than average inventories also occur in an area in the heart of the Eurasian landmass, but are most spectacularly concentrated in the north-west of North America. The languages in this latter area belong to a number of different language families with no demonstrable genealogical relationship, including Eskimo-Aleut, Na-Dene, Salishan, Tsimshianic, and Wakashan, among others. There is no evidence that the predominance of large consonant inventories in this area is a consequence of direct borrowing of words between these languages although cultural contacts between the peoples concerned are in many cases intense and deep-rooted. The situation is clearly different in one part of the African zone where large consonant inventories occur. Several Bantu languages (part of the larger Niger-Congo family) in the southern part of the continent, such as Zulu and Yeyi, are known to have enlarged their consonant inventory by borrowing clicks and other sounds which they did not previously use from languages of the Khoisan group, which already had many consonants (see, for example Louw 1975).

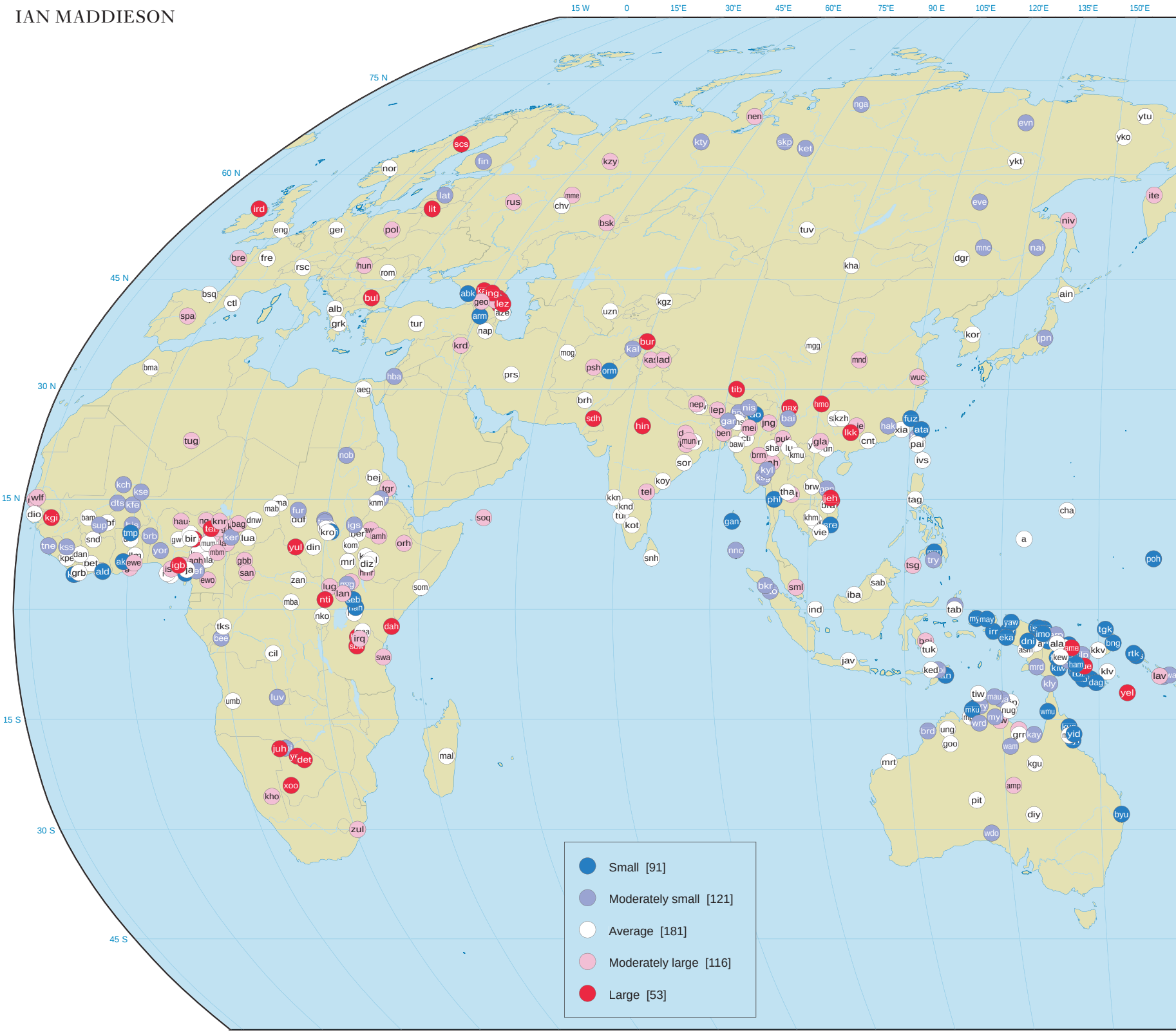
3 Theoretical issues

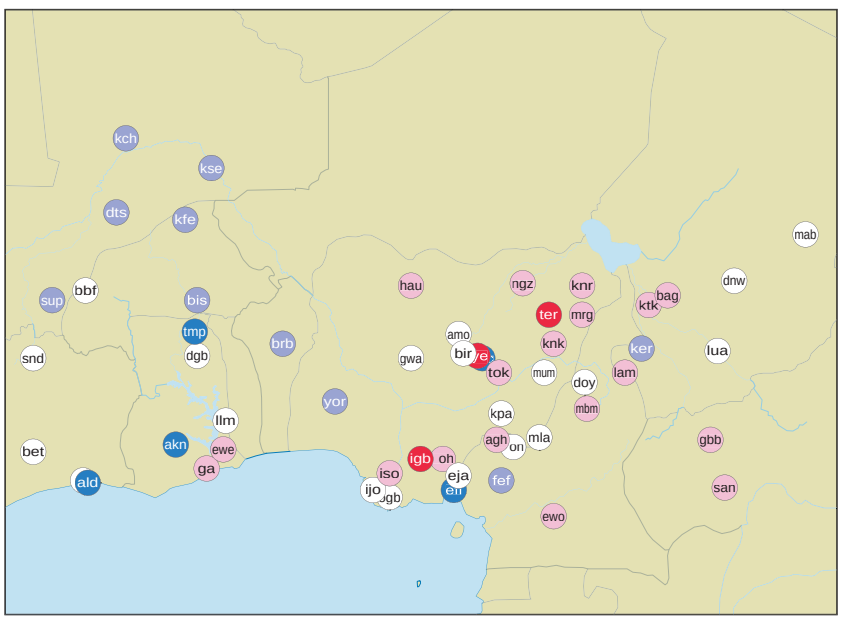
Mapping the size of consonant inventories prepares the way to investigate two connected issues. The first concerns how complexity of different aspects of the sound patterns of languages is related. All human languages are capable of expressing the range of human needs; it might therefore be assumed that they would be similar in their level of complexity. We have seen that by one simple measure of their phonological complexity, the size of the consonant inventory, languages cover quite a wide range. But complexity in one aspect might be balanced out by simplicity in another, so that in aggregate all languages are similarly complex. If this is so, mapping different aspects of phonological complexity should tend to show inverse relationships between one aspect and another in level of complexity. If this is not found, it is reasonable to conclude that languages are not constrained to be similar in this particular way, but that languages with quite different levels of complexity function just as well as each other. Several of the maps that follow will contribute to considering this question, by mapping properties of the vowel inventory (Chapter 2), the syllable structure (Chapter 12), and the presence and complexity of tone systems (Chapter 13).

The second issue concerns the hypothesis that there is an overall relationship between the size of a consonant inventory and the kind of consonants it includes. According to the "size principle" (Lindblom and Maddieson 1988) smaller consonant inventories will tend to contain only those consonants which are in various ways inherently simpler (perhaps because they involve smaller movements to pronounce them, or are easier for a listener to distinguish from other sounds). Consonants which are inherently more complex will be found in larger inventories. If this hypothesis is correct then the geographical distribution of inherently complex consonants should mirror the distribution of larger consonant inventories. In three of the following chapters, 6, 7, and 19, the occurrence of some selected classes of complex consonants will be mapped as a test of this hypothesis as well as for the inherent interest of seeing the distribution concerned.

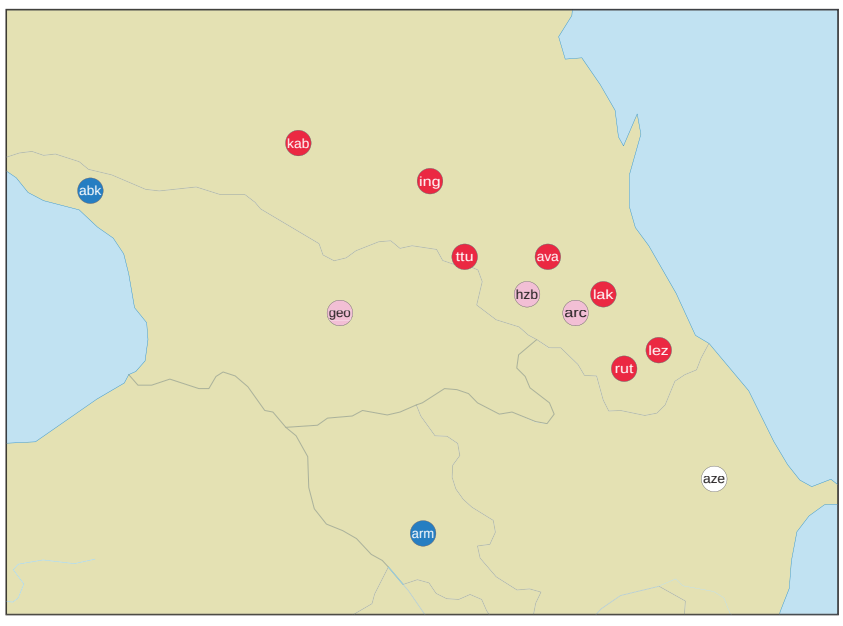
1 Consonant Inventories

IAN MADDIESON

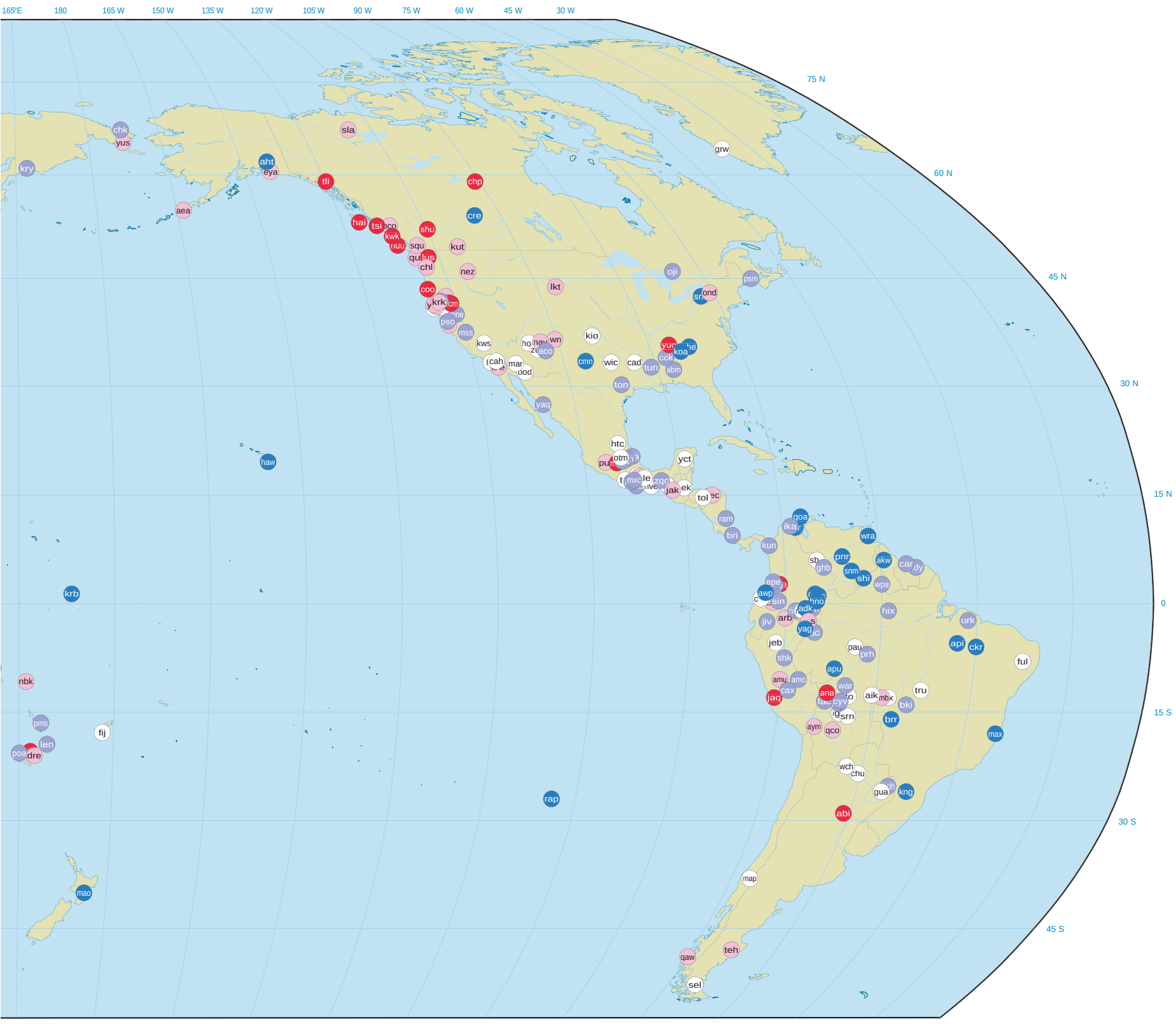


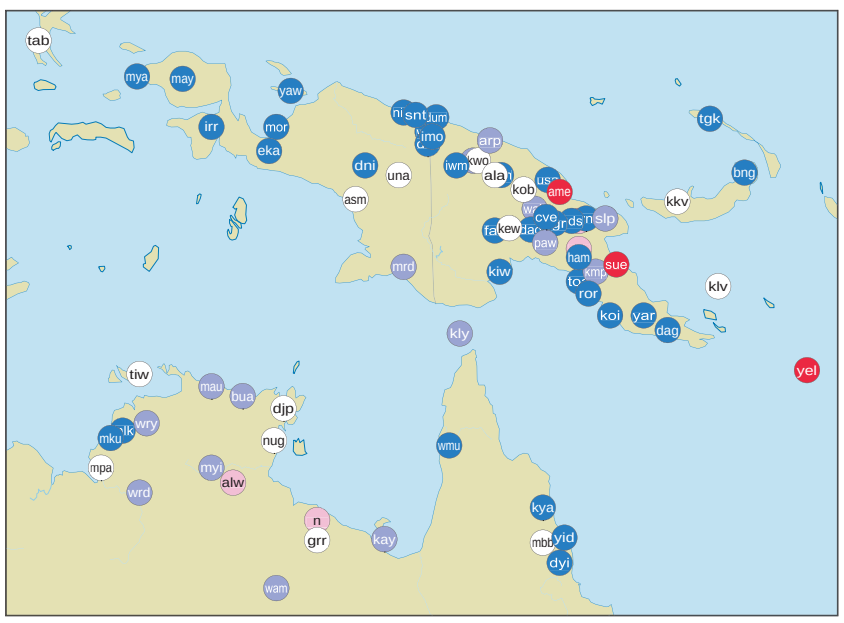


Nigeria and surrounding areas

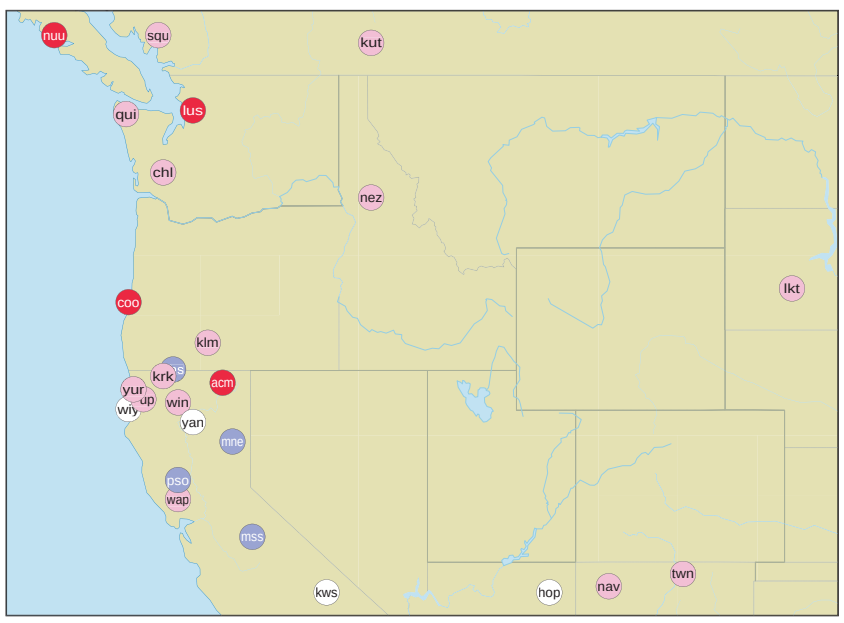


Caucasus





New Guinea and northern Australia



western United States

2 Vowel Quality Inventories

IAN MADDIESON

1 Introduction

This chapter discusses the number of vowel contrasts in the inventory of sounds in languages. It complements Chapter 1 on consonant inventories, although in this chapter the number of elements concerned is counted in a slightly different way. **Vowels** are the kinds of sounds that typically occur as the essential centres of syllables; in many languages the shortest possible word consists of just one vowel, as in the French word *eau* 'water', pronounced as the single vowel sound /o/. As this example shows, an established spelling system often involves a very indirect connection between the letters used and the vowel sounds of the language. The set of vowels used must therefore be established in the same kind of way as was discussed for consonants, in this case by comparing sets of words in which only the vowel sound differs. The English words *seat*, *sit*, *sate*, *set*, *sat*, *sot*, *soot*, *suit* show that English has at least eight different vowel sounds. By considering further sets of words with other consonants, the full set of English vowels can be established (which will vary depending on the variety of English being considered).

As with consonants, for many languages it is quite straightforward to decide how many vowels there are, but in other cases there are some difficult questions to resolve, particularly where there is a question of whether a given syllable centre should be recognized as consisting of one or of two (or even more) parts. There are a number of types of cases which pose this problem, of which only three principal ones will be discussed rather briefly here. These concern **vowel length**, **vowel nasalization**, and **diphthongs**.

In many languages a vowel which is held for a longer time contrasts with a shorter version of the same vowel (or one similar enough to be considered basically equivalent). For example, in Tlingit (Na-Dene; Alaska), the word written *t'a* /t'a/ means 'king salmon', whereas *t'aa* /t'aː/ means 'board, plank'. Such a long vowel might be considered to be two copies of the same vowel in succession, as the spelling suggests, or be thought of as a single unit. The considerations which would lead to making one choice or the other are often finely balanced and lead different scholars to different conclusions.

A nasalized vowel is one in which air is flowing out through the nose as well as through the mouth. In many languages pronunciations of the same word may vary in different contexts between saying a nasalized vowel and saying a vowel and a nasal consonant, as in the different pronunciations of French *bon* in *bon café* 'good coffee' and *bon ami* 'good friend', /bɔ̃ kafe/ versus /bɔn ami/. Again different considerations might lead to different choices, either interpreting such a case as containing a unit which is a nasalized vowel or as consisting of two parts, namely a vowel accompanied by a nasal element.

Diphthongs, such as the kinds of sounds heard between the consonants in the English words *lied* /laid/ and *loud* /laud/, involve movement of vowels from a starting position to a different finishing position. Again it is possible to consider such cases as single sounds with an inherent movement or to view them as the result of saying two different vowels in succession.

According to the decisions which are made on issues such as these the number of vowels said to occur in a given language could vary considerably. A more consistent way to compare vowel inventories is to make the comparison at a somewhat more abstract level. Phoneticians recognize three properties which contribute to the most basic quality or "timbre" of a vowel sound. These are its **height** (roughly, how open the jaw needs to be to make the vowel), its **position in a front-to-back dimension** (roughly, whether the tongue needs to be pushed forward, remain more or less in the position in which it rests during normal breathing, or be pulled toward the back of the mouth for that vowel), and the **lip position** (whether the lips are pushed forward and narrowed or not). There is much more agreement on how many vowel types differing along one or more of these basic dimensions occur in any given language. Long and short variants of the same vowel are always counted once, nasalized vowels do not add to the inventory as long as a non-nasalized counterpart occurs, and so on. For that reason, the number of basic vowel qualities in each language is what is examined in this chapter. Spanish is thus counted as having five vowels, which generally are represented quite straightforwardly by the letters *i*, *e*, *a*, *o*, *u* in Spanish orthography. The diphthongs which occur in words such as *puerta* 'door' or *siempre* 'always' can be resolved as combinations of two of these basic vowels occurring within a single syllable. For the five recognized vowel qualities in Spanish a simple phonetic categorization can be given in terms of the basic properties as follows: the two vowels represented by /i, u/ are high vowels, /e, o/ are mid vowels and /a/ is a low vowel. The vowels /i, e/ are front, /a/ is central, and /o, u/ are back; /o/ and /u/ are also rounded while the other three Spanish vowels are unrounded.

2 Establishing the values

When vowel qualities are counted in this way in the sample of languages surveyed for this chapter, the average number of vowels in a language is just fractionally below six. The smallest vowel quality inventory recorded is two and the largest fourteen. There are four languages in the sample with only two contrasting vowel qualities; these are languages in which only the height of the vowel has any distinctive function according to at least one possible interpretation of their phonetic patterns. An example of this extreme is Yimas (Lower Sepik-Ramu; Papua New Guinea). Only one language in the sample, German, uses fourteen vowel qualities and only two make use of thirteen, namely the variety of British English included here and Bété (Kru, Niger-Congo; Côte d'Ivoire). Considerably more languages have an inventory of five vowels than any other number—188 or just over one-third. The next most frequent inventory size is six vowel qualities, with 100 languages (or 17.8 per cent of the sample). In plotting the data on the map, vowel quality inventories with five or six members have therefore been grouped together in the category of "average" while those with four or fewer are classified as "small" and those with seven or more are classified as "large". Languages with "average" vowel quality inventory size account for more than half the total sample (51.2 per cent), about a third (32.5 per cent) have "large" vowel quality inventories, and only 16.3 per cent have "small" vowel quality inventories.

1. Small vowel inventory (2–4)	92
2. Average vowel inventory (5–6)	288
3. Large vowel inventory (7–14)	183
total	563

3 Geographical distribution

There are strong areal patterns in the distribution of vowel quality inventories. Not surprisingly, languages with average inventory sizes are the most widely scattered. In just a few areas, southern Africa being one, they occur almost to the exclusion of the other two types. Small and large inventories on the other hand are markedly skewed in their geographical distributions. Languages with small inventories are frequent in the Americas. The indigenous languages of the Americas quite often have four vowels in a set similar to that of Spanish except for missing an /u/-like vowel. Examples of languages having this kind of system are Eastern Ojibwa (Algonquian;

Ontario), Navajo (Athapaskan; southwestern United States), North Puebla Nahuatl (Uto-Aztecan; Mexico), and Tacana (Tacanan; Bolivia). Others of these languages, such as Aymara, Cherokee, and Haida, have three-vowel systems, usually reported as having two high vowels /i, u/ and a low central vowel /a/. In Australia small vowel inventories dominate, this feature being just one of a number of properties which tend to give the languages native to this part of the world a special character from the point of view of their sound patterns. Small vowel inventories occur rarely in the remaining parts of the world, that is, in Africa, the entire Eurasian mainland, and New Guinea and the Pacific Islands, although there are some specific small language groups, such as the Berber languages of North Africa and the North-West Caucasian languages spoken near the border of Russia and Georgia, which may have this feature.

Africa is strikingly marked by a zone right across the "middle belt", roughly between the Equator and the Sahara, in which large vowel inventories predominate. This belt encompasses languages belonging to three major families, Niger-Congo, Nilo-Saharan, and Afro-Asiatic. The relatively large number of vowels in these languages seems to be associated with the prevalence of patterns of **vowel harmony** in the same area. When a language is said to have vowel harmony this generally means that within a word, including any affixes, it is only possible to combine the members of certain subsets of the vowels together. Such restrictions are very common in both Niger-Congo and Nilo-Saharan languages within this area. Since its effect is to reduce the number of legitimate words that can be constructed from the set of sounds available, the presence of vowel harmony may make it easier to tolerate a larger than average number of different vowels in a language as it reduces the risk of mishearing and misidentifying a word. This is because in a language with vowel harmony a given vowel in a word does not have to be discriminated from among the set of all possible vowels, but only among those of a subset.

There are also concentrations of larger than average vowel inventories in interior South-East Asia and southern China, in much of Europe and, on a smaller scale, in interior New Guinea. The European area also includes a number of languages with vowel harmony restrictions on the distribution of their vowels, such as Finnish and Hungarian, as well as Turkish and its relatives (which extend well into Central Asia and to western China). However, large vowel inventories in some of the other languages in this area came about (in part, at least) as a result of earlier distinctions between sets of long and short vowels being transmuted into contrasts of vowel quality. This occurred (subject to other influences as well) in English, German, and Italian, amongst others.

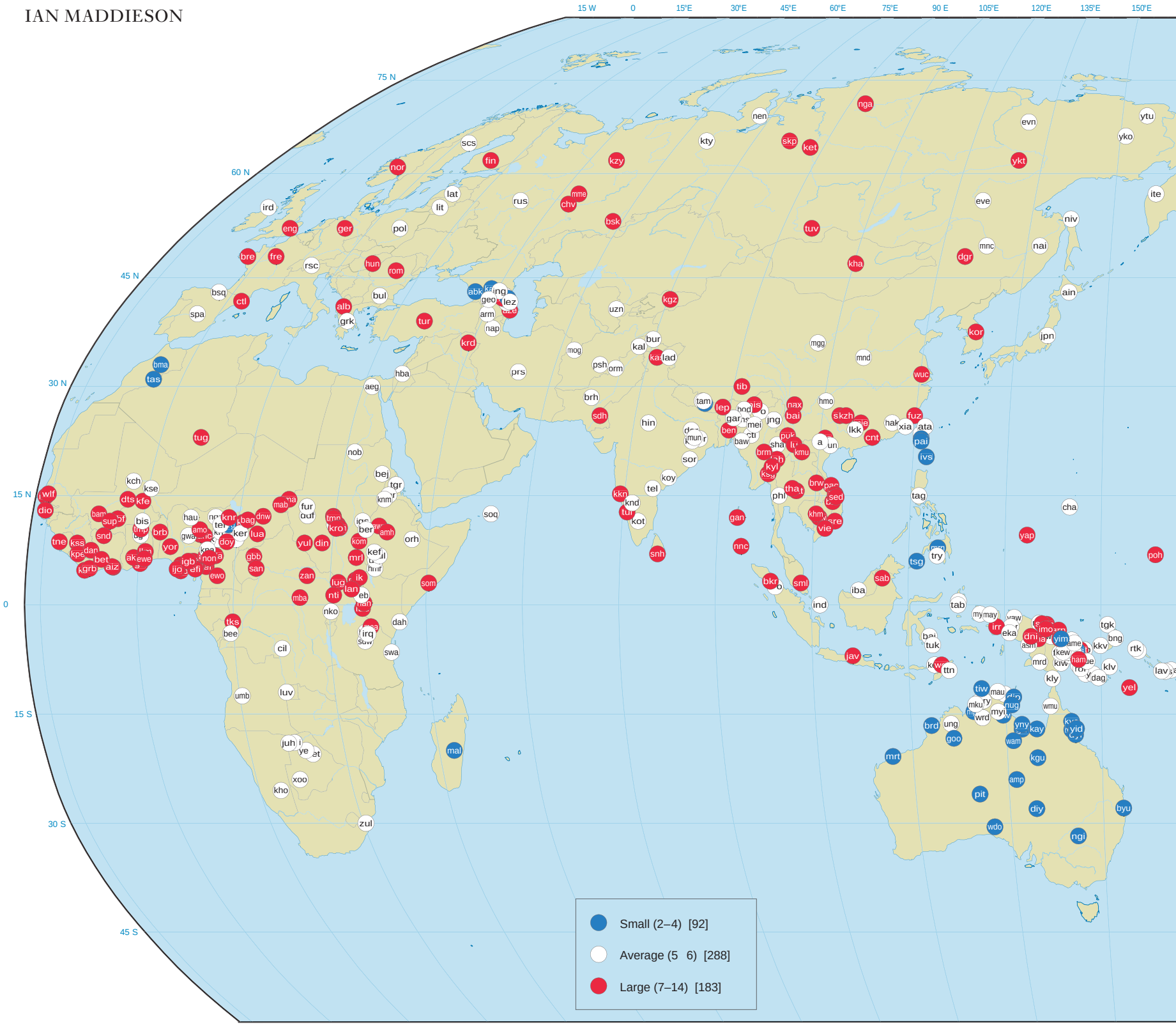
4 Discussion

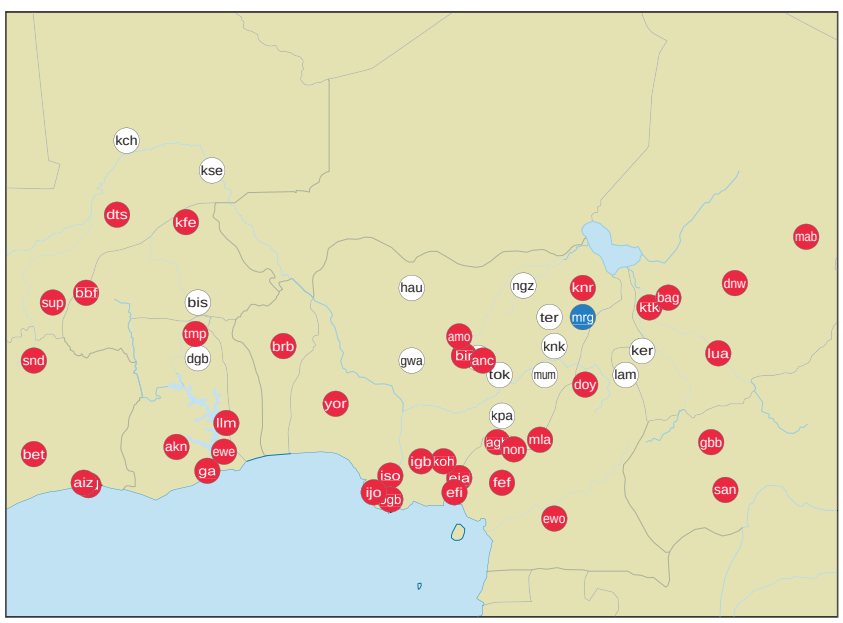
In the Asian area concerned a number of the languages are well known for having undergone historical sound changes which shortened their words so that many of them became just one syllable long, and often at the same time reduced the number of distinctions between different consonants, especially at the ends of words. In some cases at least, these changes also resulted in increasing the number of distinct vowels. Such patterns of interrelated changes have sometimes been taken to suggest that languages maintain an overall balance in the complexity of their phonological systems. It is therefore interesting to examine whether a tendency to balance a small number of consonants with a larger number of vowels and conversely to balance a larger number of consonants with fewer vowel contrasts can be detected in the data of this survey. In a set of 559 languages for which the consonant inventory size and the vowel quality inventory size are both available, absolutely no correlation was found between the number of vowels and the number of consonants (for the statistically minded, the correlation coefficient between the two series of numbers is −.004). The lack of any significant relationship between these properties is also evident when the broad size categories discussed in this chapter and Chapter 1 are considered. For example, of the 169 languages with "moderately large" and "large" consonant inventories as defined in Chapter 1, 33 have "small" vowel quality inventories and 50 have "large" vowel quality inventories, or 19.8 per cent and 29.9 per cent respectively. These proportions are very similar to the overall frequencies found in the entire sample. As a general principle it cannot therefore be said that the numbers of vowels and consonants in an inventory, as counted here, bear a relation to each other. There are many different possible balances between these two aspects of a language's sound system which all serve to create a satisfactory tool for human use.

The relationship between vowel inventory size and consonant inventory size is further discussed from a different perspective in Chapter 3, and the relationship between vowel systems and tone systems is discussed in Chapter 13.

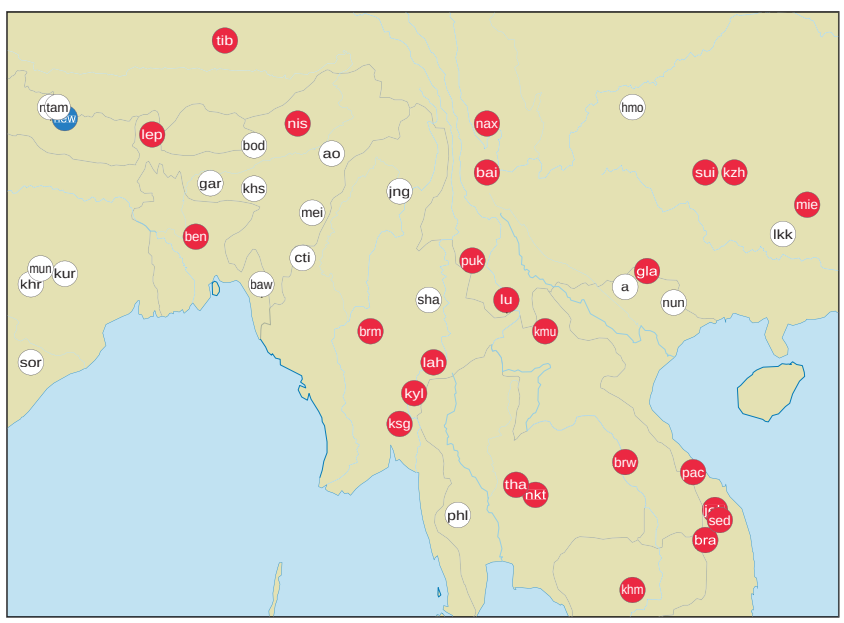
2 Vowel Quality Inventories

IAN MADDIESON

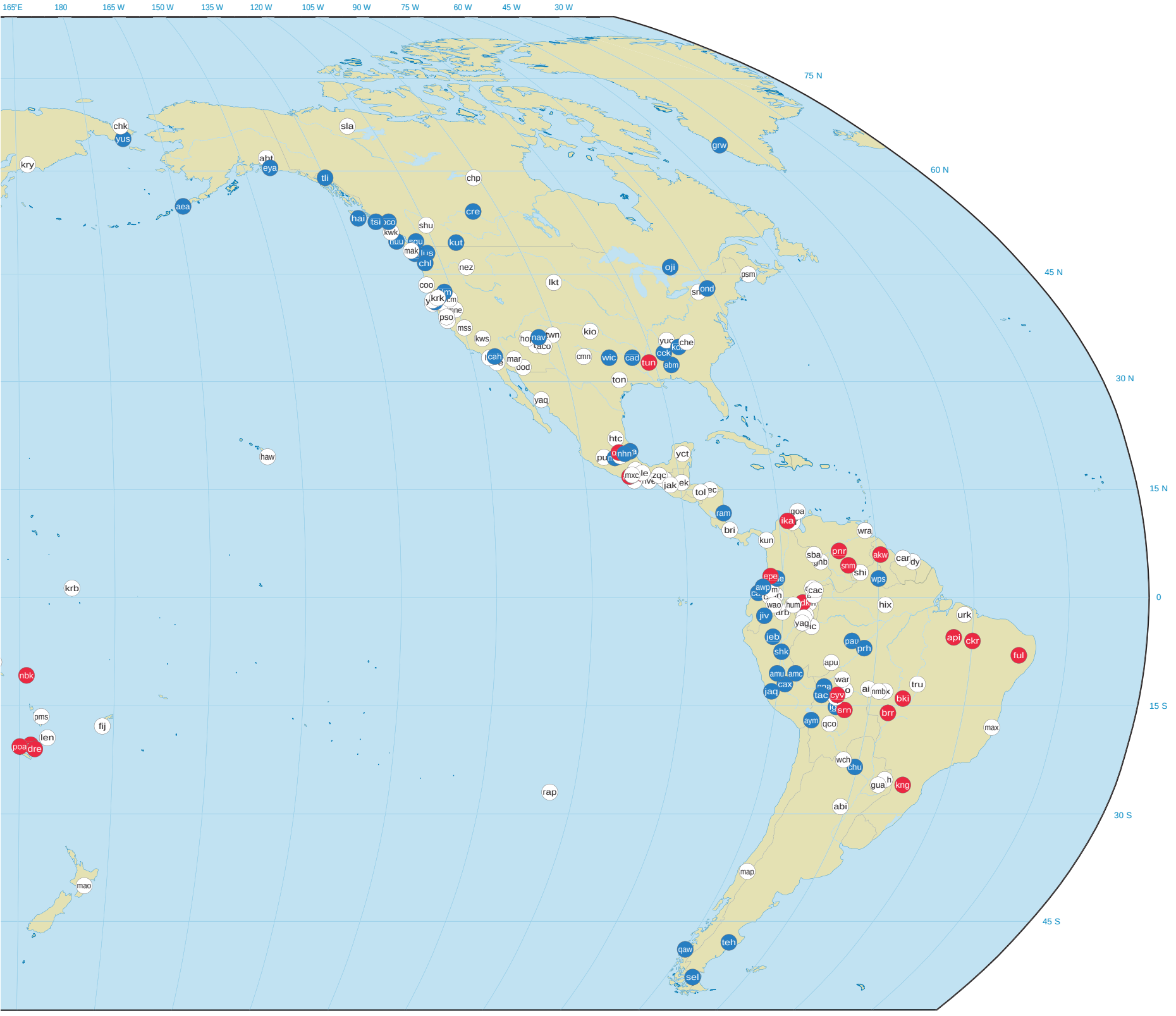


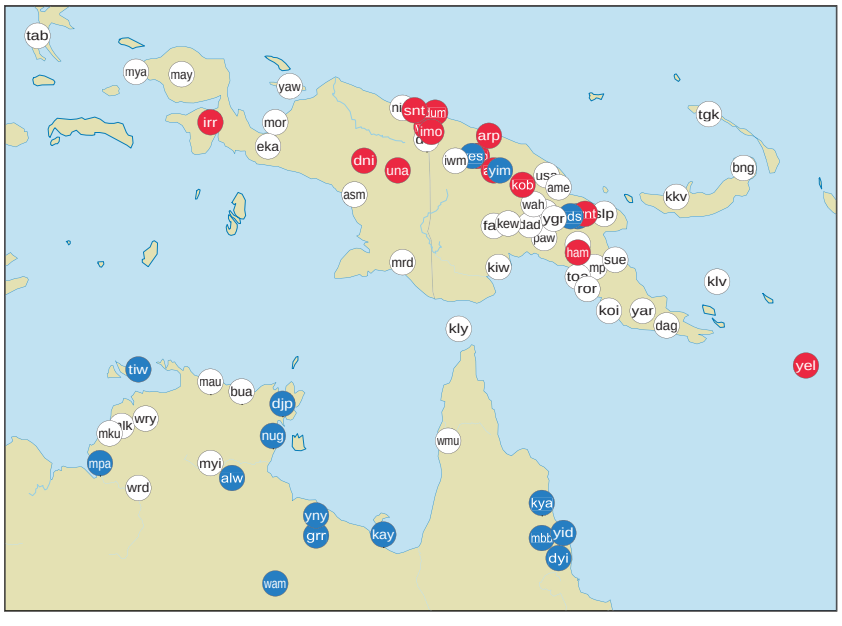


Nigeria and surrounding areas

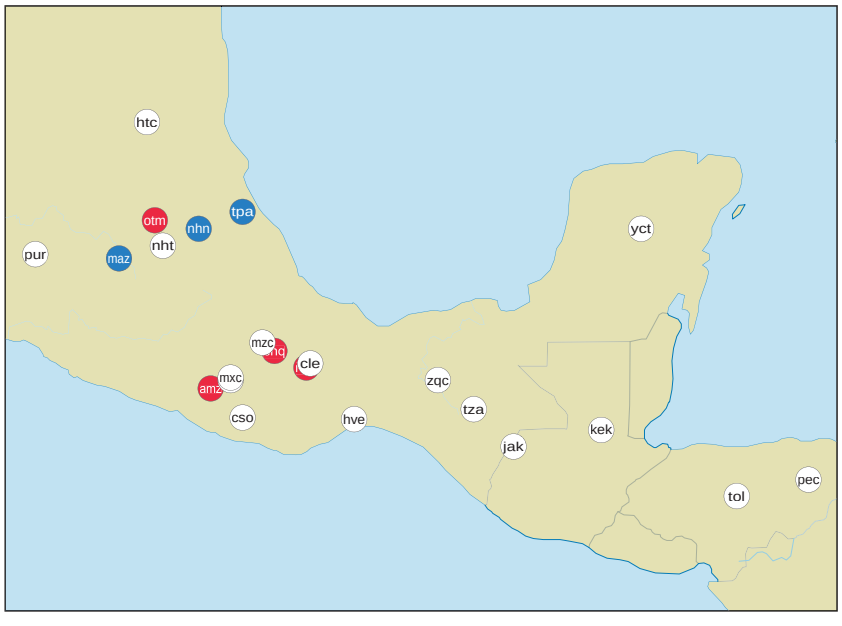


Myanmar and surrounding areas





New Guinea and northern Australia



Mesoamerica

3 Consonant–Vowel Ratio

IAN MADDIESON

1 Introduction

In previous chapters the size of the consonant inventory (Chapter 1) and the number of distinct vowel qualities (Chapter 2) have been discussed separately. In this chapter the ratio between these two properties of the sound systems of languages around the world will be considered. This is one way in which the complexity of different aspects of the systems can be evaluated. In Chapter 2 it was noted that there is no overall correlation between consonant inventory size and the number of vowel qualities. By looking at the two aspects jointly a more subtle picture of their interaction can be obtained.

The ratio is calculated simply by dividing the number of consonants (C) by the number of vowel qualities (VQ) and will be referred to as the C/VQ ratio. The resulting numbers range from a low of only a little over one to a high of twenty-nine. The lowest value among the 563 languages for which it has been calculated is represented by Andoke (isolate; Colombia), which has ten consonants and nine vowel qualities. The highest number is represented by Abkhaz (North-West Caucasian; Georgia), which is analysed as having fifty-eight consonants but only two vowel qualities. The ratio thus ranges between 1.11 and 29, but the more common values are nearer the lower end of the range: the mean is 4.25 and the median 3.5. The languages were grouped into five categories based on dividing the range into convenient steps below, near to, and above the median so as to create a histogram with approximately a normal distribution. Languages with a ratio equal to or below 2.0 were classed as having a "low" C/VQ ratio. Those with a ratio above 2.0 but below 2.75 were classed as "moderately low". Those with a ratio of 2.75 or higher but less than 4.5 were classed as "average", those with values of 4.5 or higher but less than 6.5 as "moderately high", and those with values above 6.5 as "high". Only ten languages have ratios of twelve or higher.

1. Low	59
2. Moderately low	97
3. Average	234
4. Moderately high	102
5. High	71
total	563

2 Geographical distribution

The most common category shown on the accompanying map represents languages with an average C/VQ ratio. These languages—ones with roughly three or four times as many consonants as vowel qualities in their sound inventories—are particularly common in West Africa and are quite common in New Guinea and island Asia, South-East Asia, Central America, and the eastern side of South America, as well as in northern Eurasia. With the exception of the last area it is striking that these languages are mostly clustered relatively close to the equator.

The 156 languages with proportionately low numbers of consonants are shown on the map as the two categories having a "moderately low" C/VQ ratio or a "low" ratio. There are no particularly noticeable areas which are characterized by concentrated numbers of languages from only one of these categories, but there is a very strong overlap between the distribution of these two categories and that of the "average" category. Apart from a cluster mostly of Chadic languages the remainder of the languages of West Africa fall into one of these two classes. Similarly most of the other languages in the sample from New Guinea and island Asia, much of South-East Asia as well as the eastern side of South America have "low" or "moderately low" ratios.

On the other hand, the 173 languages with "moderately high" or "high" C/VQ ratios show marked clustering. Languages with higher than average ratios predominate in Australia, in most of eastern and southern Africa, and along the western side of the Americas. There is a marked cluster of languages with "high" C/VQ ratio in the Caucasus, and another in southern Africa, but the most substantial group of languages of this class occurs in the north-west of North America. This is an area that is also distinctive with respect to many other measures of phonological typology, particularly ones relating to consonant systems (see, for example, Chapters 7 and 8).

3 Discussion

These observations suggest that the most important distinction to make from the geographical point of view with respect to C/VQ ratios is that between those languages which are average or below and those which are above the average ratio. Southern Africa, the Caucasus, and the North American North-West Coast are characterized by large consonant inventories, and indeed most of the

Table 1 Co-occurrence of vowel quality and consonant inventory sizes

	Consonant Inventory Size				
VQ size	*small*	*mod. small*	*average*	*mod. large*	*large*
small	11	19	27	22	11
average	51	71	80	53	31
large	29	31	73	40	10

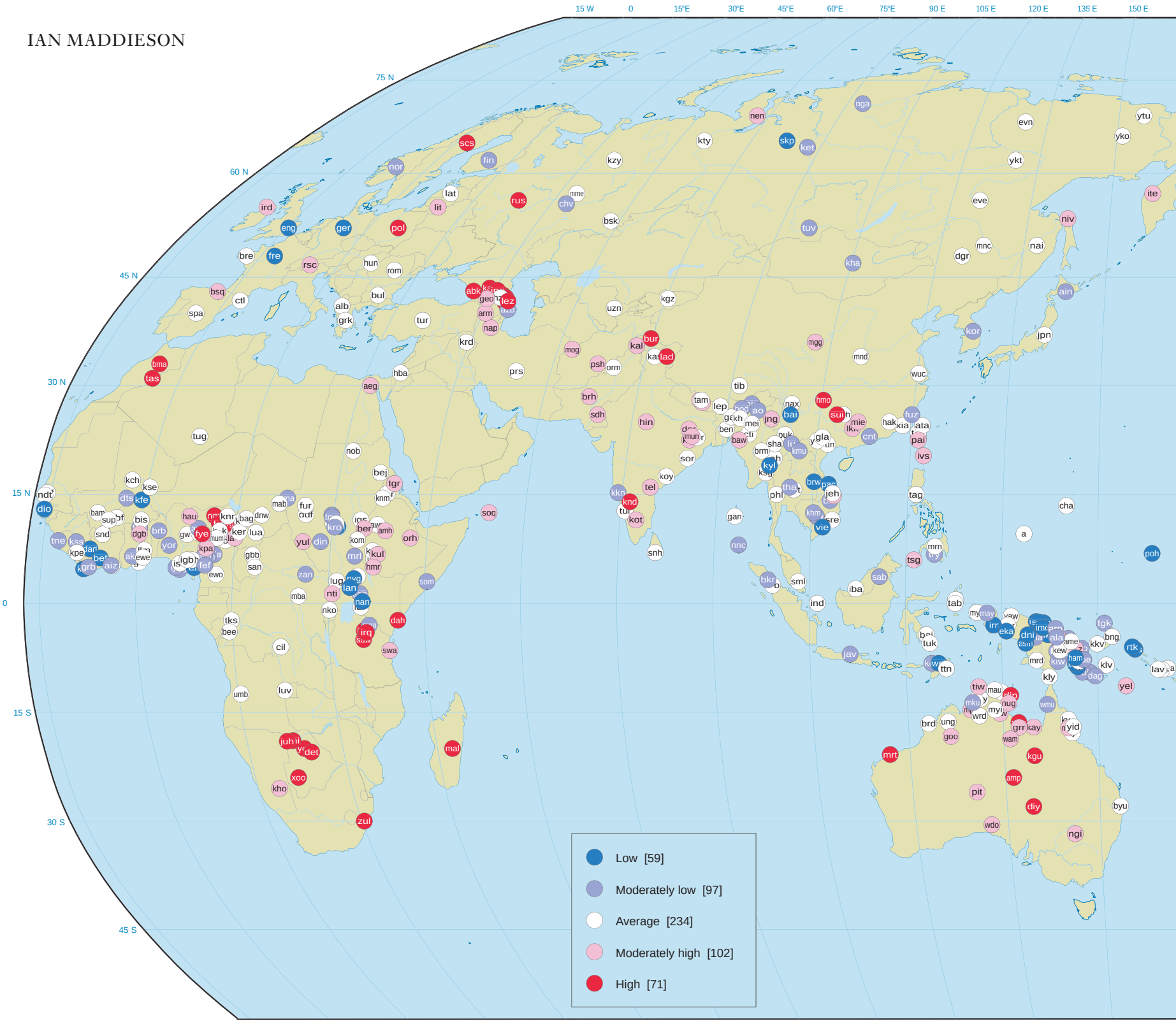
languages in the sample with large consonant inventories also have high C/VQ ratios and none has a C/VQ ratio below average. The facts which produce this result have probably led to the suggestion that it is a typical pattern for large consonant inventories to be counterbalanced by smaller vowel systems across the languages of the world. However, in the West African area, larger than average consonant systems occur together with larger than average vowel systems, leading these languages to mostly have average or below-average C/VQ ratios, and in Australia the languages typically combine small vowel inventories with smaller than average consonant inventories, resulting in a higher than average C/VQ ratio. Using the categories established in Chapters 1 and 2 gives the contingency table shown as Table 1. A sample of 559 languages is represented in the table.

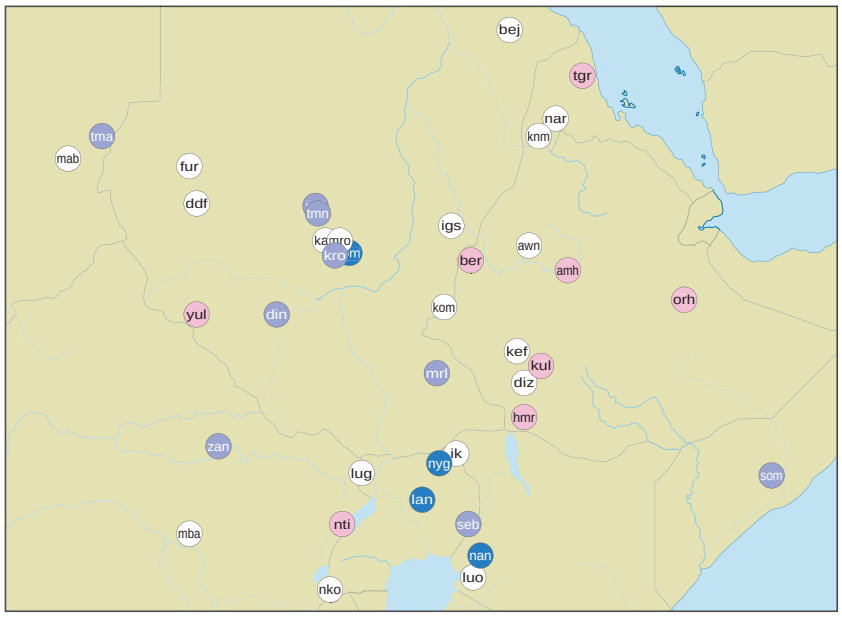
There are 11 languages with small consonant inventories and a small number of vowel qualities, and 11 combining a large consonant inventory and a small number of vowel qualities. There are 10 combining a large inventory of both vowels and consonants. A not very much larger number, 29, combine a large number of vowel qualities with a small inventory of consonants.

From these numbers it is apparent that the occurrence of a large consonant inventory with a small number of vowel distinctions is not part of a general pattern in languages but reflects a geographically restricted tendency that can be found in a few areas (primarily in southern Africa, the Caucasus, and the American north-west). Moreover, there are other ways for a language to be 'consonant-rich', as in the case of Australian languages such as Diyari (Pama-Nyungan; South Australia), analysed as having a very average 22 consonants but only three vowel qualities, giving a high C/VQ ratio of 7.33.

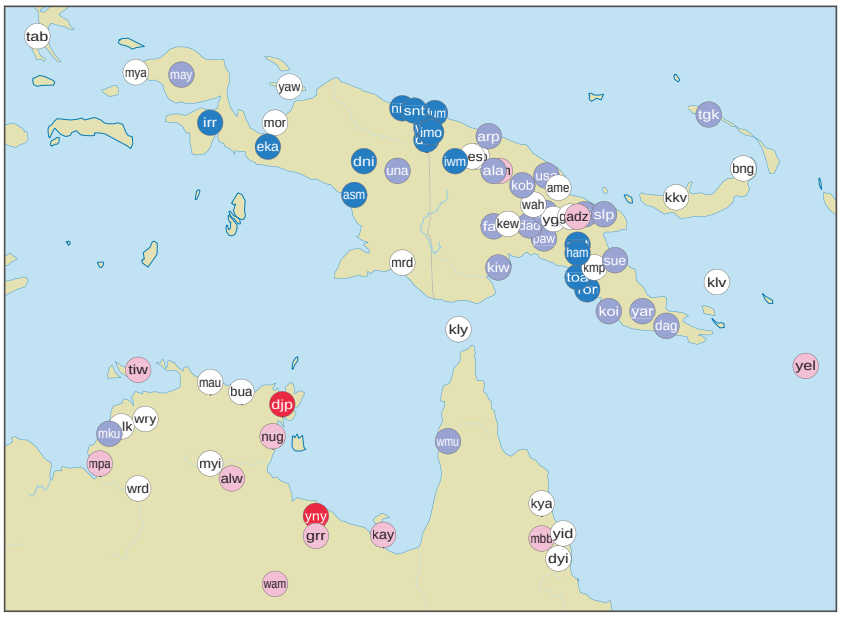
3 Consonant–Vowel Ratio

IAN MADDIESON

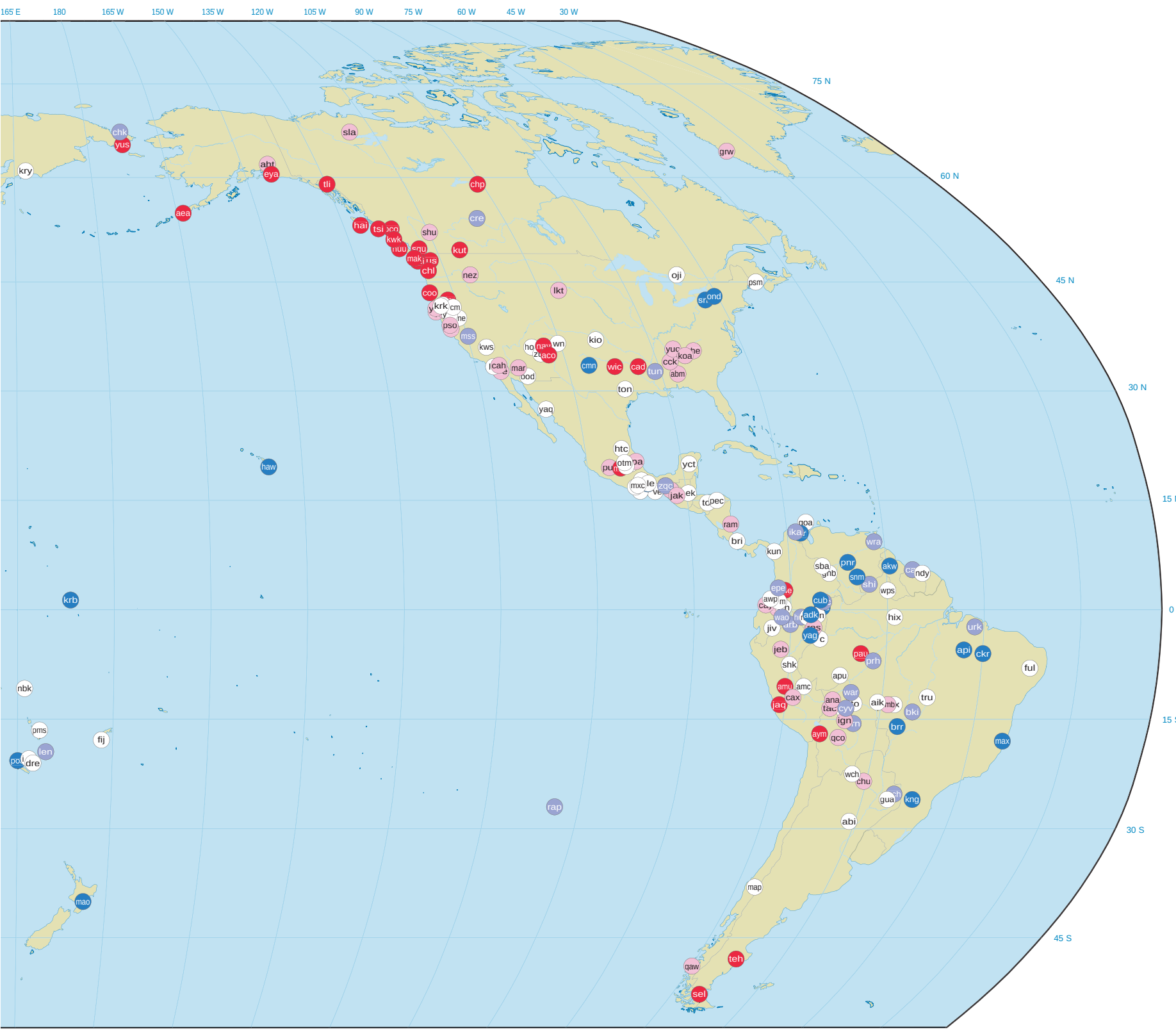


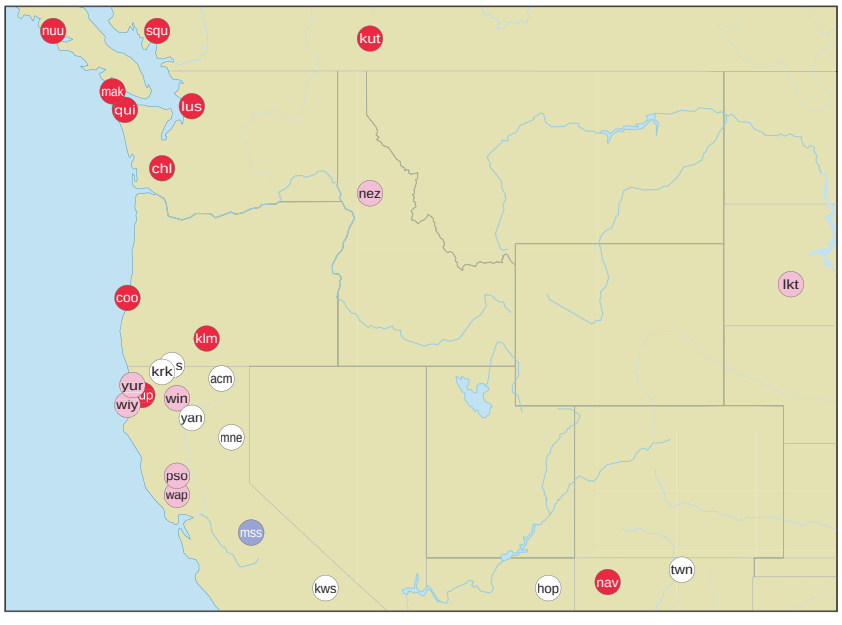


eastern Africa

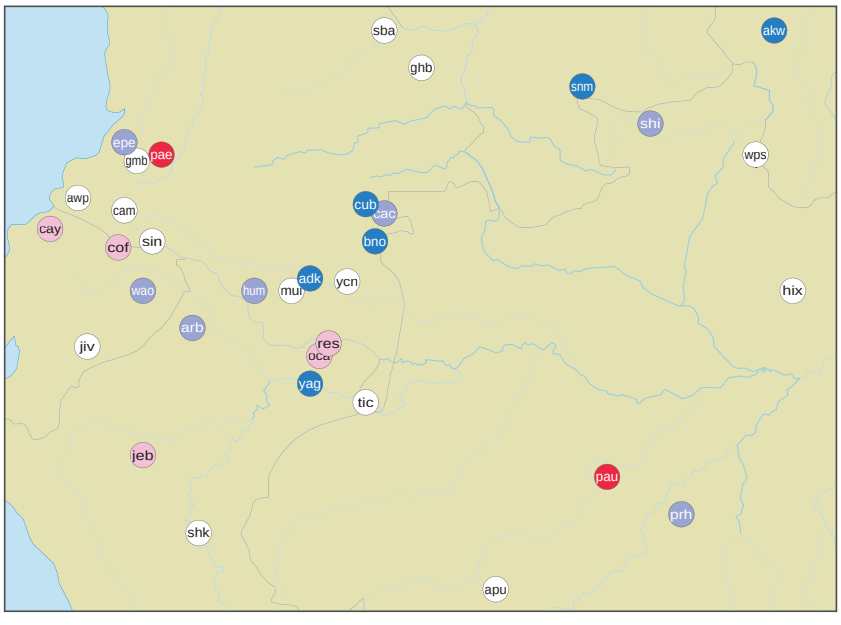


New Guinea and northern Australia





western North America



northwestern South America

4 Voicing in Plosives and Fricatives

IAN MADDIESON

1 Introduction

As noted in Chapter 5, one of the major ways that consonants differ from each other is in the accompanying action of the larynx, with the most typical larynx settings being one which allows air to flow freely between the vocal folds versus one in which the vocal folds vibrate to produce regular voicing. One aspect of this opposition between voiced and voiceless consonants is discussed in Chapter 5. In this chapter the distribution of a contrast between voiced and voiceless counterparts in the two major classes of consonants within which this contrast commonly occurs will be discussed. These two classes of consonants are the plosives and fricatives. **Plosives** are the kinds of sounds usually associated with the letters *p*, *t*, *k*; *b*, *d*, *g*, in which airflow from the lungs is interrupted by a complete closure being made in the mouth. **Fricatives** are the kinds of sounds usually associated with letters such as *f*, *s*; *v*, *z*, in which the air passes through a narrow constriction that causes the air to flow turbulently and thus create a noisy sound. The other classes of consonants which are found in the majority of languages (nasals, "liquids", and vowel-like approximants) are voiced in the overwhelming majority of cases.

Voiceless plosive and fricative consonants occur in more languages than voiced ones, but voiced types are nonetheless relatively common. How frequent it is for a contrast between voiced and voiceless plosives and fricatives to occur in languages and how such contrasts are distributed will be the focus of this chapter. A language will only be counted as having a contrast between voiced and voiceless plosives or fricatives if there is a pair of sounds in which the place of articulation and all other principal characteristics of the pair apart from the voicing category are the same. Thus, for English a pair of words such as *rope* and *robe* are sufficient to establish that English has contrastive voicing in plosives, since the final consonants in these two words are both bilabial plosives and they differ in one being voiceless and the other voiced. Similarly, the pair of words *rice* and *rise* are sufficient to show that English has a voicing contrast in fricatives, since the final consonants in these two words are both alveolar fricatives which differ in voicing. English in fact has three pairs of contrasting plosives and four pairs of contrasting fricatives. The language Chickasaw (Muskogean; Alabama and Mississippi) has three voiceless plosives (bilabial, dental, and velar in place), but has only one voiced one, the bilabial, written /b/. This, however, is sufficient for this language to be counted as having a voicing contrast in plosives. On the other hand Seneca (Iroquoian; New York State) is reported as having the two voiceless plosives /t, k/, which are articulated at alveolar and velar places, and a voiced bilabial plosive, /b/. This language is not counted as having a contrast of voicing in its plosives since there is no voiced/voiceless pair at any one place of articulation. Canela-Krahô (Ge-Kaingang; Brazil) is reported to have one voiceless fricative, which is velar in place, and two voiced fricatives, written /v, z/, which are pronounced at the labiodental and alveolar places. Therefore this language is not counted as having a voicing contrast in fricatives, since there is no pair of voiced and voiceless fricatives made at the same place.

○	1. No voicing contrast	181
●	2. Voicing contrast in plosives alone	189
●	3. Voicing contrast in fricatives alone	38
●	4. Voicing contrast in both plosives and fricatives	158
	total	566

2 Geographical distribution of voicing contrasts

About a third of the languages surveyed for this chapter (32.0 per cent) have **no voicing contrast in either plosives or fricatives**, as this has been defined above. The largest concentration of these occurs in Australia, where the great majority of the languages have only voiceless plosives and lack fricatives altogether (see Chapter 18). A substantial number of languages in the Americas also fall into this class but in these cases there are usually both plosives and fricatives present in the consonant inventories. One typical example is Zuni (isolate; New Mexico), which has several voiceless members of both the plosive and fricative consonant classes but no voiced members of either of these classes. Canela-Krahô, mentioned above, is a less typical case since it has both voiced and voiceless fricatives but nonetheless lacks contrast. Some of the languages in Africa and the Pacific islands which fall into this class have both voiced and voiceless plosives but the voiced members of the class are prenasalized, that is, they begin with a part during which air is flowing out through the nasal passage. An example is Paamese (Oceanic; Vanuatu). This has the voiceless plosives /p, t, k/ but their voiced counterparts are only heard with a nasal onset as part of the prenasalized voiced plosives, so there is no straightforward contrast of voicing in plosives in this language. Absence of any voicing contrast also occurs sporadically in East Asia.

Approximately another third (33.4 per cent) of the languages surveyed have **a voicing contrast in plosives but not in fricatives**. This type is dominant in the most southerly parts of Asia, where it is typical of the Dravidian languages as well as languages from other families, and is prevalent in New Guinea. It is also well represented in Africa and the Americas, but is largely absent from Europe and western Asia. These languages most typically have some fricatives but only voiceless ones. An African example is Yoruba (Defoid, Niger-Congo; southeastern Nigeria), which has three contrasting pairs of voiced and voiceless plosives, as well as three voiceless fricatives. Chickasaw, mentioned earlier, also belongs in this group although it has only the one contrasting plosive pair. It has four voiceless fricatives but no voiced ones. Ika (Chibchan; Colombia) and Murle (Surmic, Nilo-Saharan; Ethiopia) are unusual members of the group since they are reported as having only voiced fricatives, which therefore do not contrast with voiceless counterparts. Huave (Huavean; Oaxaca, Mexico) is reported to have some voiceless and some voiced fricatives, but no pairs at the same place of articulation, so this language also belongs in this group.

Another substantial group of languages has **a voiced/voiceless contrast in both plosives and fricatives**. This group amounts to 28.1 per cent of the total, so it has slightly fewer members than the two previously described. Languages of this group are dominant in Europe and western Asia and also very common in Africa, but are quite rare outside these areas of the Old World. Most of the principal languages of European colonialism such as Portuguese, English, French, and Russian, now widely used outside their original homelands, belong to this group, and their influence on indigenous languages in areas of their spread is likely to "recruit" progressively more languages into the group (if the indigenous languages are not simply replaced). Note that Spanish is not treated as having a voicing contrast in plosives since the sounds written with the letters *b*, *d*, *g* are not pronounced as plosives in most of their occurrences in speech but as voiced fricatives or approximants. Spanish therefore belongs to the final group of languages in this classification, those with a voicing contrast in fricatives but not in plosives.

Only a relatively small proportion of the languages, 6.7 per cent of the total, have **a voicing contrast in fricatives but not in plosives**. These languages do not show much tendency to cluster geographically except for a small group in the western subarctic zone, where this pattern is found in the Eskimo-Aleut languages as well as in some of the adjacent Na-Dene languages. Siberian Yupik (Eskimo-Aleut; Siberia), for example, has three contrasting voiced/voiceless pairs of fricatives, but no voiced plosives. Other

examples of this group include Tsou, UMbundu, and Mesa Grande Diegueño. Tsou (Austronesian; Taiwan) has the contrasting fricative pairs /f, v/ and /s, z/ but only /p, t, k/ in its set of plosives. Tsou does, however, have the implosives /ɓ, ɗ/, discussed in Chapter 7. UMbundu (Bantu; Angola-Namibia border) has the fricative pair /f, v/ in contrast but the only voiced stops in the language are prenasalized plosives. Mesa Grande Diegueño (Yuman; California) has five voiceless fricatives and six voiceless plosives, but the only direct voicing contrast it has comes in the pair of lateral fricatives (see Chapter 8).

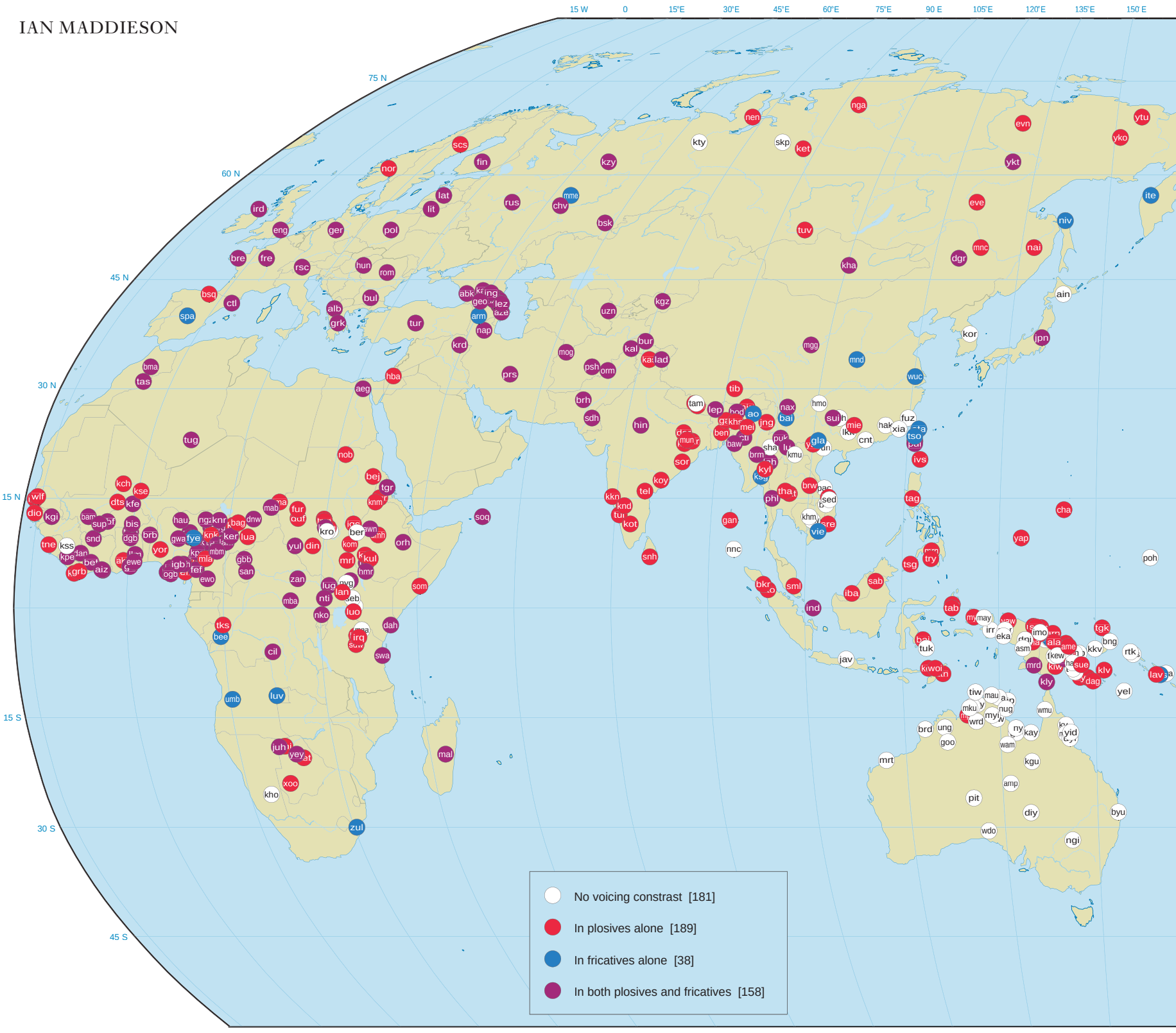
3 Discussion

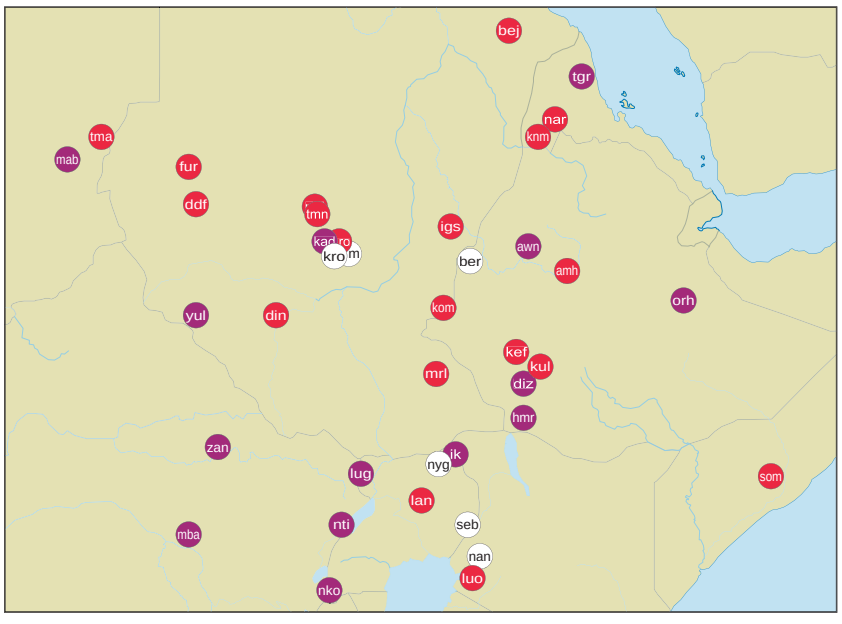
When numbers are pooled across the four categories of languages discussed in the chapter, we see that a voicing contrast in plosives is considerably more common than a voicing contrast in fricatives. In all, 347 or 61.3 per cent of the languages in the survey have a voiced/voiceless contrast between at least one pair of plosives either with or without any contrast in voicing among fricatives. Only 196 or 34.5 per cent have a voicing contrast between at least one pair of fricatives, either with or without any contrast in voicing among plosives.

Since there are so many more languages in which a fricative voicing contrast occurs in combination with a plosive voicing contrast than those in which it occurs alone, this suggests that there is some preference for constructing languages in a way that makes fricative voicing "parasitic" on plosive voicing. We can use the overall frequencies to calculate what numbers would be expected in our sample if plosive voicing and fricative voicing were independently distributed. These calculations show that about 120 languages would be expected to have the combination of the two voicing contrasts, and about 78 would be expected to have a fricative voicing contrast alone. These numbers are sufficiently different from the observed totals of 158 and 38 to show that the occurrence of fricative voicing is significantly dependent on the presence of plosive voicing in the same language. At the same time, however, the occurrence of a fricative voicing contrast does show a decided geographical limitation, being especially an Old World phenomenon, and very largely absent from the Americas, Australia, and New Guinea, nor even particularly frequent in South and East Asia. Co-occurrence of voicing contrasts in both plosives and fricatives might therefore be largely due to an areal spread of fricative voicing in the languages of the Old World.

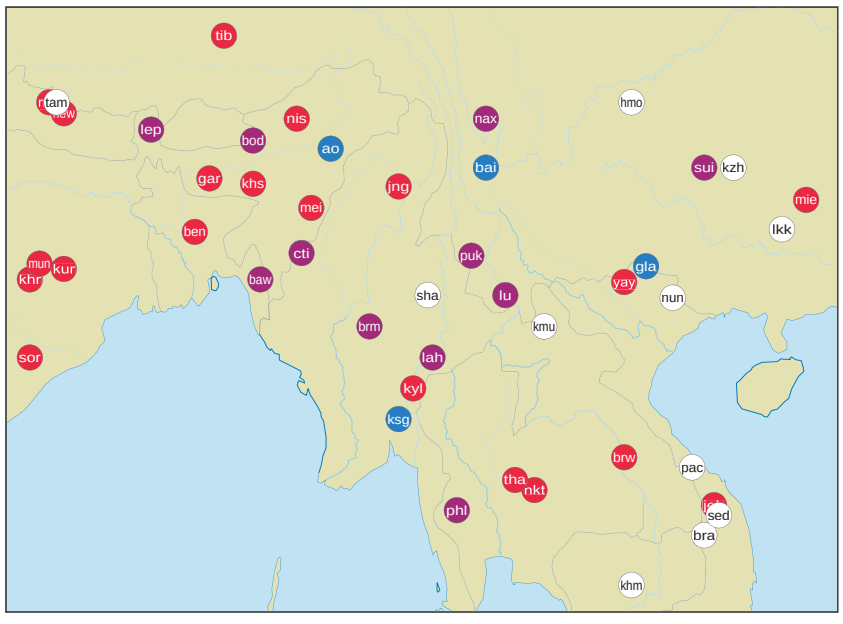
4 Voicing in Plosives and Fricatives

IAN MADDIESON

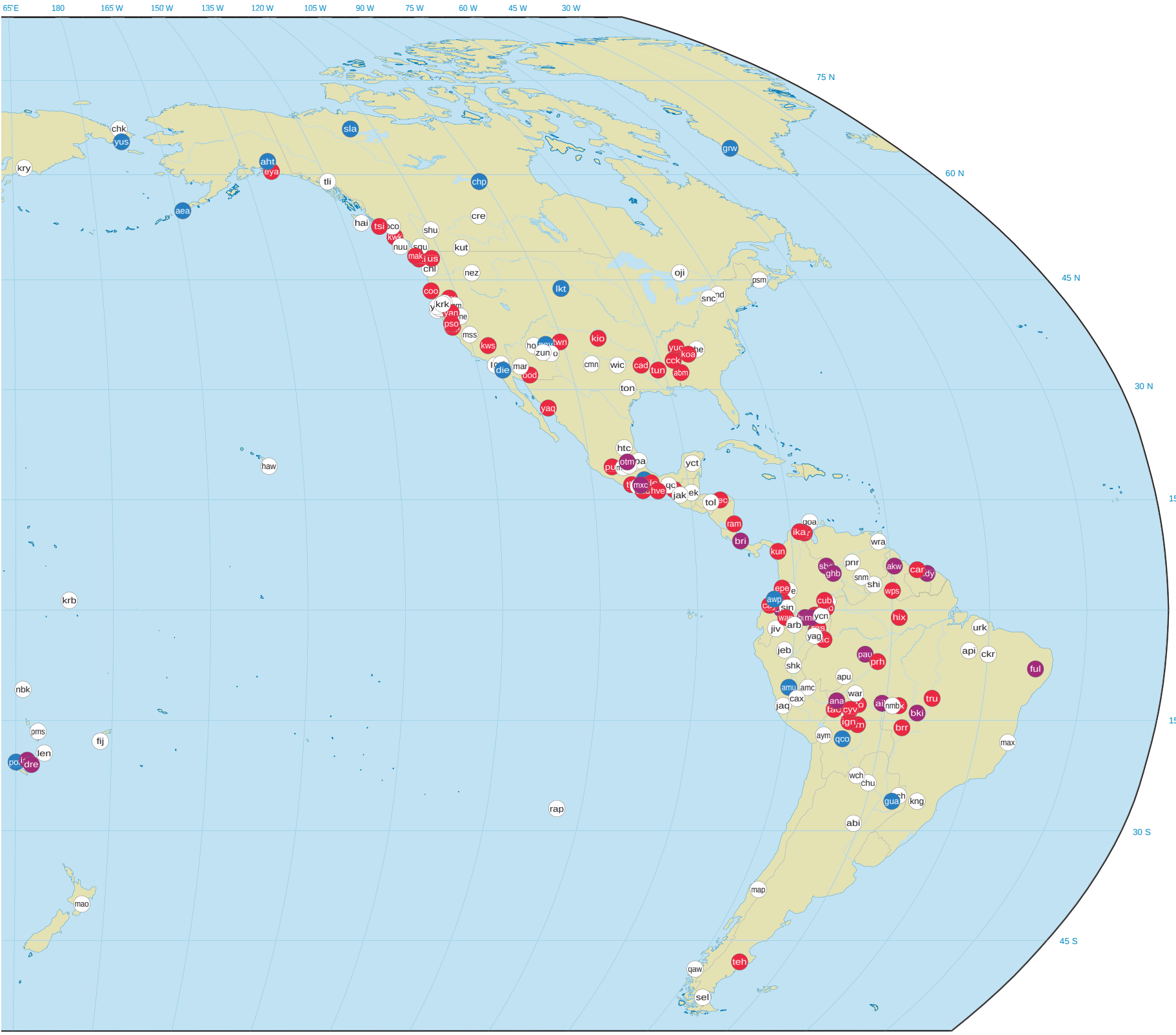


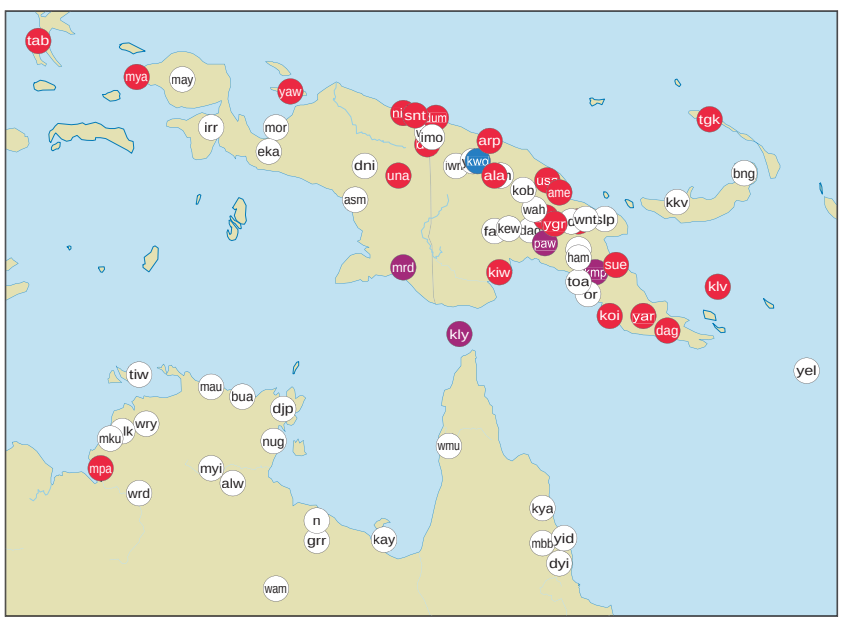


eastern Africa

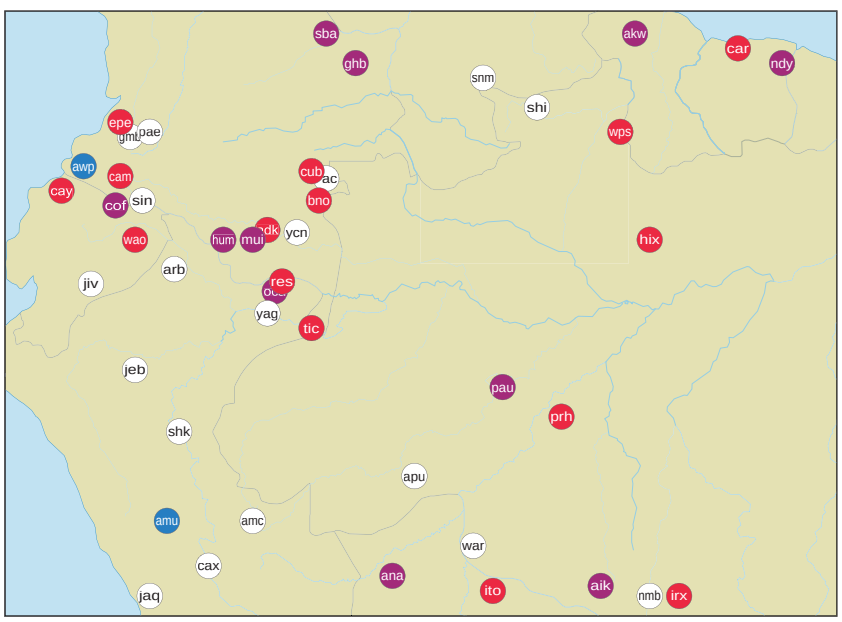


Myanmar and surrounding areas





New Guinea and northern Australia



northwestern South America

5 Voicing and Gaps in Plosive Systems

IAN MADDIESON

1 Plosive systems

In this chapter we will look at some patterns found within the sets of stop consonant sounds in the world's languages. It will emerge that two well-known patterns, here labelled "missing /p/" and "missing /g/", which were previously considered to reflect universal phonetic factors in an equal way, are quite differently distributed. This difference suggests that these patterns are not in fact parallel.

1.	Other	243
2.	/p t k b d g/	256
3.	Missing /p/	32
4.	Missing /g/	32
5.	Both missing	3
	total	566

Stops are sounds in which the flow of air which is active in creating the sound is completely blocked for a short interval of time. In the most common type of stop sound, known as a **plosive**, air in the lungs is briefly blocked from flowing out through the mouth and nose, and pressure builds up behind the blockage. The sounds that are generally associated with the letters *p*, *t*, *k*, *b*, *d*, *g* in English words such *pat*, *kid*, *bag* are examples of plosives. At least some sounds of this type are found in every known language spoken by humans. In phonetic terminology the sounds associated with *p*, *b* are called **bilabial plosives** because their pronunciation requires bringing the two lips together; the sounds associated with *t*, *d* are called **dental** or **alveolar plosives** because the tongue closes against the upper teeth or the skin covering the roots (alveoli) of the teeth; the sounds associated with *k, g* are called **velar plosives** because the tongue closes against the soft part of the palate, technically known as the velum. Bilabial, dental, alveolar, and velar are known as **places of articulation**.

The letters *p, t, k* represent **voiceless plosives**, that is, those in which the vocal folds in the larynx are apart and air is able to flow quite freely from the lungs into the mouth cavity. The letters *b*, *d*, *g* represent **voiced plosives**, that is, those in which the vocal folds are in a position close together so that, provided the air pressure in the lungs is sufficient, the vocal folds will be caused to vibrate against one another (as they do, for example during humming). Almost half of the languages surveyed for this chapter include these six plosives. In some languages, these are the only stops. In many others there are additional types of stops, such as plosives made by closing the tongue against the uvula as described in Chapter 6, or stops made with other configurations of the larynx, such as the glottalized stops discussed in Chapter 7, or the clicks mentioned in Chapter 19, to mention only a few of the possibilities. None of these other stops will be taken into account in this chapter.

2 The values

2.1 Other. The first class of languages shown on Map 5 are those which neither include all six of the set /p, t, k, b, d, g/ (now using these letters as general phonetic symbols) in their consonant inventories, nor have one of the three specific variants on this set which will be described shortly. These languages have a system of stop sounds which is simply described as "other" than those of specific interest here. In all, 243 (or 43 per cent of the sample) languages fall into the category. There are many different ways in which a language can qualify as a member of the "other" category. The most common case, accounting for about 17 per cent of the total sample, is that the language lacks voiced stops of any kind, as in Mandarin, Nahuatl, Dyirbal (Pama-Nyungan; Queensland, Australia), or Maori. But there are many other possibilities; for example, the language may have voiced stops but not of the simple plosive type, as is the case in Fijian or Vietnamese, or it may lack bilabial consonants of any type, as is the case with Tlingit (Na-Dene; Alaska), or it may have partial sets of voiced and voiceless plosives, as in Seneca (Iroquoian; New York State), which has only three of the six plosives under discussion, namely /t, k, b/. And these are far from the only patterns. As this class is very heterogeneous it is of somewhat limited interest to examine the geographical distribution it displays, but including all such cases makes it possible to provide an estimate of the overall frequency of the patterns that are coded by the remaining categories.

2.2 All six stops /p t k b d g/. Of the remaining languages, 256 (or 45 per cent of the sample) have all six of the stops under discussion, with or without additional stops of one or more kinds. This class is designated "/ptkbdg/". Languages of this type predominate in the greater part of the Eurasian landmass and in sub-Saharan Africa, and are not uncommon in the Americas and in New Guinea. They are almost entirely absent from Australia, from the southern cone of South America, from the northern tier of North America, and from a part of East Asia. These are areas where the languages mostly lack any simple voiced plosives in their consonant inventories and hence do not have voiced counterparts to /p, t, k/.

2.3 /p/, /g/ or both missing. The residual languages have one or more gaps of specific interest in the set of plosives they contain. It was noted by linguists some time ago that if one of the common set of voiceless plosives is absent from a language it is more likely to be /p/ than either /t/ or /k/; and if one of the common set of voiced plosives is absent from a language it is more likely to be /g/ than either /b/ or /d/. These absences might be explained by how the flow of air from the lungs during speaking interacts with the movements of the speech articulators as they are positioned to make different kinds of sounds. In a plosive the regular outflow of air is briefly held back by the closure in the mouth. After this closure is formed, the pressure of the air in the mouth cavity quickly reaches the level of the pressure which is driving air out from the lungs. If the vocal folds are in the position for voicing, this will happen more slowly, since the rate of airflow from the lungs is slowed down by the narrowed passage in the larynx. However, since the vocal-fold vibration which we call voicing is driven by the flow of air between the vocal folds, voicing will not be able to continue when the air pressure above the larynx approaches that below the larynx, as the flow will become insufficient to drive the vibration, which consequently will stop. How rapidly this happens is related to how large the space in the mouth is between the larynx and the location of the plosive closure. It will take longest in /b/, since the closure is as far away as possible from the larynx and the enclosed space is the largest possible, and, importantly, the possibilities for expansion of this space by yielding of the soft tissues of the cheeks and other surfaces under pressure is greatest. By contrast, in /g/ the space is much smaller because the location of the closure is much closer to the larynx. Consequently, voicing is more likely to be extinguished before the plosive closure is released when the pronunciation target is /g/ than when it is /b/. This could possibly lead to confusion of /g/ with /k/, and over time to loss of the distinction between the two sounds. Alternatively, if the plosive is pronounced with less than a complete closure (as often occurs in more relaxed speech) and consequently voicing is able to continue through its duration, the pronunciation norm may shift away from the plosive realization. Either path may provide a route by which /g/ is eliminated from the

consonant set. Finally, in a language which is undergoing a process creating a series of voiced plosives, the difficulty of combining voicing with velar articulation may prevent such a process from effecting a change of /k/ into /g/ under the same conditions which change /p/ into /b/ and /t/ into /d/. Because of the aerodynamic facts, /g/ can be seen to be a less favoured plosive than /b/ or /d/. (For a more detailed discussion of these issues, see Ohala 1983.)

It has sometimes been considered that bilabial plosives are least compatible with voicelessness for a reason also linked to speech aerodynamics, namely, the large volume of the cavity behind the closure at the lips allows the greatest volume of air to flow between the open vocal folds from the lungs into the mouth cavity before that cavity becomes fully pressurized and the flow stops. For this reason, the amount of pressure built up behind a closure at the lips might be less at the time the closure is released than is the case with a closure formed further back in the mouth, especially if the closure is only held for a short time. Consequently the force of the explosion of air at the release of a voiceless bilabial plosive may be weaker than for other plosives, making it more likely to be misidentified. In fact, detailed phonetic studies show that almost all of the pressure increase in the mouth during a plosive occurs in the first 20 milliseconds of the closure duration, and furthermore that the closure for /p/ tends to be held a little longer than that for /t/ or /k/. Hence differences in pressure in the mouth are unlikely to be relevant to explaining why /p/ seems less favoured than /t/ or /k/.

The residual categories on the map show the distribution of languages whose plosive sets are lacking a /g/ sound or a /p/ sound when it might have been expected to occur. Stringent criteria were applied in order to make sure that the expectation is well founded; there must be a distinction between voiced and voiceless plosives, and the relevant place of articulation must be present in an otherwise standard-looking plosive set. Thus in the pattern that is labelled "missing /p/" the language has /b/ in its series of voiced plosives and lacks /p/ in its series of voiceless plosives but otherwise has matching pairs of voiced and voiceless plosives, as in (1). In the pattern labelled "missing /g/" the language has /k/ in its series of voiceless plosives and lacks /g/ in its series of voiced plosives but otherwise has matching pairs of voiced and voiceless plosives, as in (2). A few languages are labelled "both missing"; they include the sounds shown in (3) in their plosive set.

(1)				(2)				(3)			
		t	k		p	t	k			t	k
	b	d	g		b	d			b	d	

There are equal numbers of cases of "missing /p/" and "missing /g/" in the sample of languages examined, namely 32 of each. In addition there are three languages which show the "both missing" pattern: Una (Trans-New Guinea), Ket (Yeniseian; Siberia), and Efik (Cross River, Niger-Congo; southeastern Nigeria). In terms of frequency of occurrence, then, these two patterns are very similar. However, the map shows that they are not geographically distributed in similar fashion.

3 Geographical distribution

Languages representing the "missing /g/" pattern are quite widely scattered in the Americas, and are also found in Africa, in New Guinea and island Asia, and on the Asian mainland. Although a local concentration occurs in South-East Asia, the impression is essentially of a widespread global distribution—precisely what would be expected if this pattern is one which may arise in multiple independent localities in response to general physical laws governing speech production.

On the other hand, languages representing the "missing /p/" pattern occur with a much more clustered distribution, with a predominant cluster in Africa north of the equator, especially around the margins of the Sahara desert, in languages from all three of the major language families present in the area (Niger-Congo, Nilo-Saharan, and Afro-Asiatic). Only a few cases occur outside this area, most notably a small group in New Guinea. The African concentration weakens the assumption that there are similarly general principles working against the combination of bilabial place and voicelessness as is the case for the combination of velar place and voicing—note that the proposed aerodynamic explanation for "missing /p/" appears invalid. Instead, it suggests that one or more areal factors may have operated to favour the initiation and/or spread of this pattern within the circum-Saharan zone.

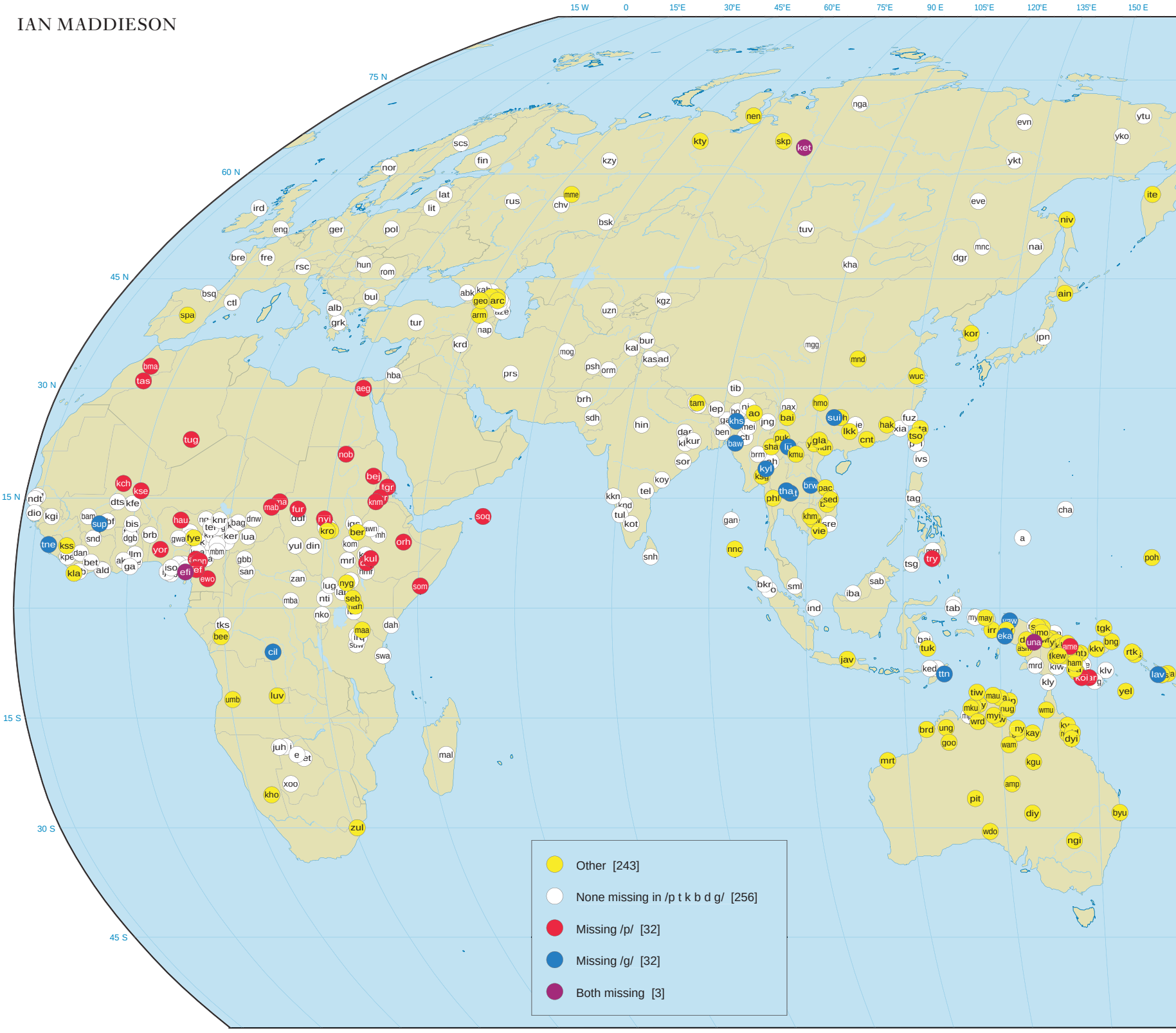
These areal factors may have been as subtle as the convergence on a regional norm for pronunciation of /p/-sounds. Imagine that voiceless plosives came to be typically pronounced with aspiration, that is, with the voiceless airflow continuing for some time after the closure has been released, as if pronouncing /h/ after /p, t, k/. Because when a stop closure is released the lips separate more quickly than happens with the other articulators, the noise generated by the release itself is shorter and less easy to distinguish from the noise of the following aspiration in /ph/ than is the case with /th/ or /kh/. Hence only the sound of /h/, or a sound reminiscent of a labial fricative such as /f/, might be identified by a listener (compare Ancient Greek /ph/ as the source of /f/ in modern words such as 'phonetics'), and the pronunciation norm would again change.

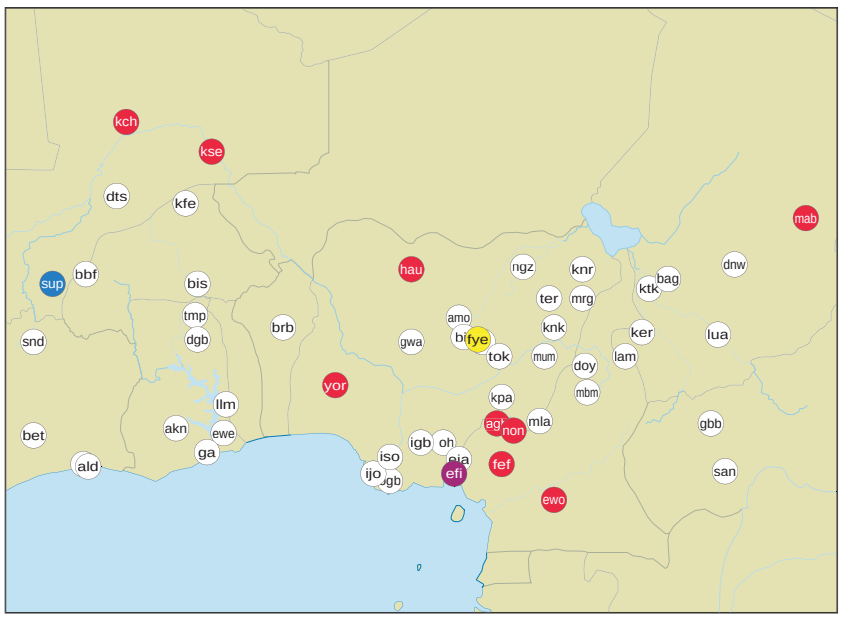
It is also possible to speculate on roles for more straighforwardly social factors, such as the prestige of Arabic—a "missing /p/" language—over much of this area for centuries, or the potential impact of cosmetic modification of the lips, practised in historic times among quite a number of the peoples of the Sahel, of the northern margins of the tropical rain forest, and into the Ethiopian highlands.

A more detailed description of plosive systems and the gaps in their structures would take into account matters such as where these sounds occur within the structure of words, the relative frequency of different plosives in the lexicon, or their relation to other sounds not considered here. However, the narrow perspective of the present analysis is sufficient to suggest that a useful distinction can be drawn between the very general factors tending to exclude /g/ from consonant inventories, and the more locale- or language-specific ones tending to exclude /p/ from consonant inventories.

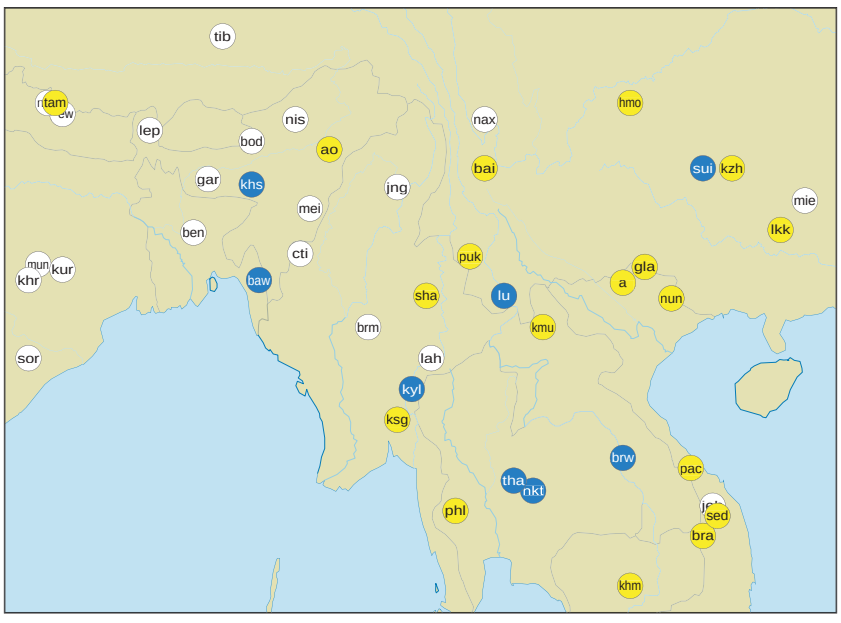
5 Voicing and Gaps in Plosive Systems

IAN MADDIESON

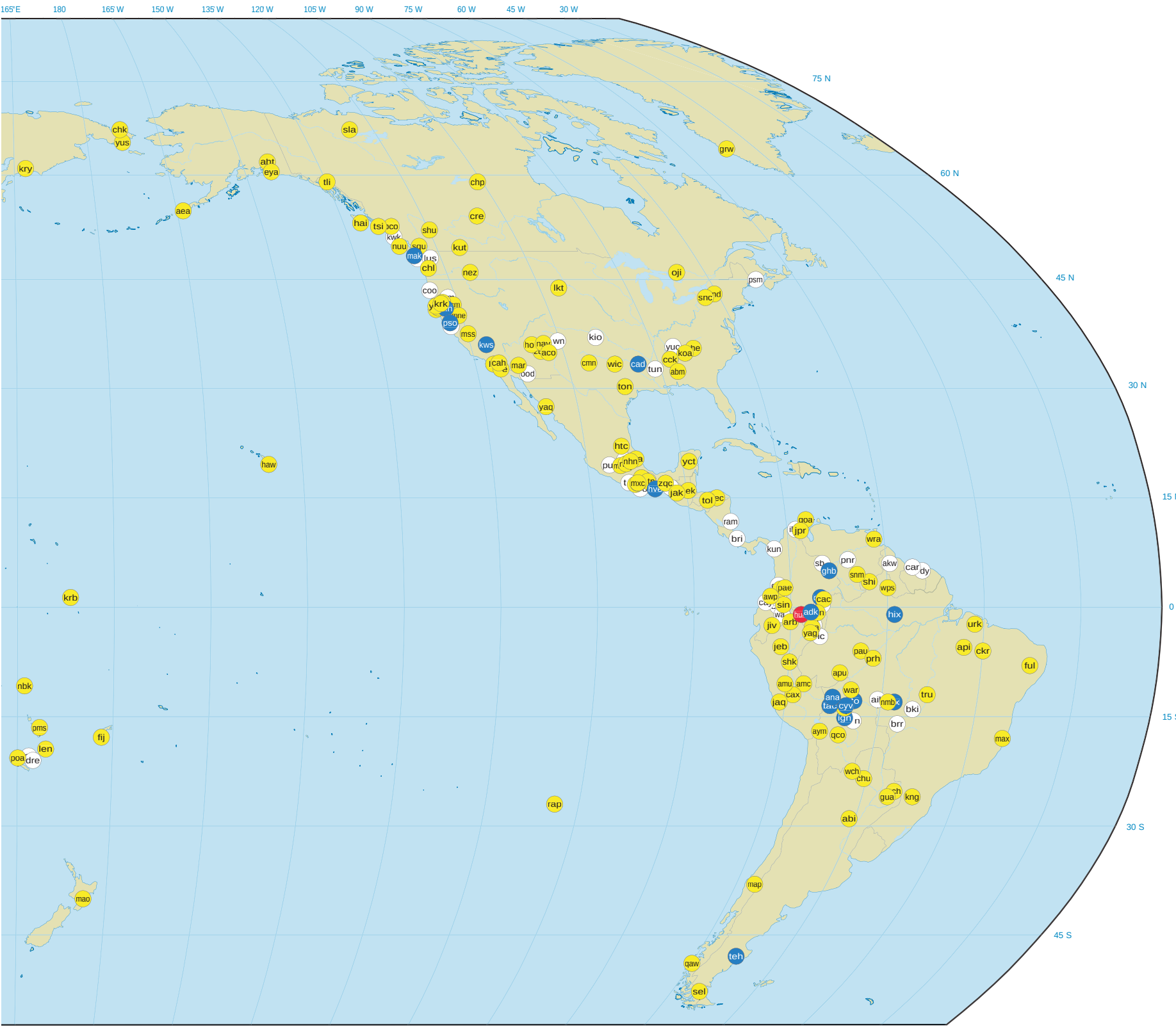




Nigeria and surrounding areas



Myanmar and surrounding areas



New Guinea and northern Australia

northwestern South America

6 Uvular Consonants

IAN MADDIESON

1 Introduction

This chapter will discuss the distribution of one of the less common types of consonants, those with a uvular place of articulation. Consonants of this class are not as rare as those discussed in Chapter 19, but are nonetheless found in only roughly one sixth of the languages surveyed. Most languages have consonants in which the back of the tongue is raised toward the velum; some also have consonants in which the back of the tongue is raised toward the point where the opening into the nasal cavity is located. These are called **uvular consonants**.

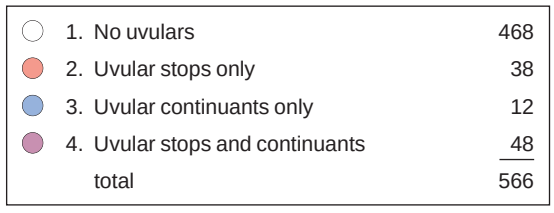
Toward the back of the roof of the mouth the soft part of the palate narrows and terminates in a fleshy "tail" which is the uvula. *Uvular* is the adjective derived from *uvula*. Besides its role in speech, the uvula is active in actions such as snoring and gargling and plays a part in regulating the connecting passage between the mouth and the nasal cavity. When air is flowing in or out through the nose the back part of the soft palate (or velum) together with the uvula is lowered to open the passage between the nasal cavity and the mouth. The sound of loud snoring is caused by the uvula vibrating as air is forced through a narrower opening between mouth and nose than occurs during normal breathing. Most of the time during speaking this opening is closed by the velum being raised and the uvula making contact with the back wall of the pharynx so that the flow of air through the nose is blocked.

2 Defining the values

The great majority of the languages surveyed (80.9 per cent) have no uvulars. Uvulars are absent from several large areas, such as the northern part of South America, the eastern part of North America, West and Central Africa, southern Asia including Indonesia and the Philippines, and the Pacific region except for a couple of languages in New Guinea. They are also absent from most of the languages of Europe outside the Caucasus region. Uvulars are concentrated in the western part of North America and the southern part of South America, in the Caucasus and Inner Asia as well as the far north-east of Asia. A lesser concentration occurs in eastern and southern Africa. The languages concerned are grouped into different classes depending on whether the uvulars found are stops, continuants, or both, as detailed below.

The most common type of uvular consonant is a stop of one kind or another, including plosives, affricates, and ejective stops (ejectives are defined in Chapter 7). In phonetic transcription voiceless and voiced uvular plosives are written with the symbols [q] and [ɢ] respectively, and a uvular ejective stop as [q']. In 86 of the languages surveyed for this chapter—almost nine-tenths of those with any uvulars—there are one or more uvular stops. In 38 of these languages the only consonants at the uvular place of articulation are stops. These languages form the first group in our classification.

The other uvular consonants are continuants of various kinds. These are consonants during whose production the flow of air is continuous, unlike stops, which the airflow is briefly held back by a closure in the mouth. Continuants include fricatives and nasals, as well as trills and approximants. Of the languages surveyed, 60 have uvular continuants of one kind or another. In 48 of these this is in addition to stops, and in 12 the only uvular consonants are continuants. Languages with only continuants among their uvulars are grouped into a second category, and those with both uvular stops and continuants form the third.



3 Uvular stops and continuants

Languages whose only uvulars are stops are found especially in Africa and South America, but occur also in all areas where uvulars occur. In almost all the languages with uvular stops they contrast with velar stops, as, for example, in the Coast Tsimshian (British Columbia) words /q'ool/ 'six' versus /k'ool/ 'one' (these number words are used with reference to specific classes of objects). Only a very few languages, such as Qawasqar (Alacalufan; Chile), have uvular stops—as in the name of the people and their language—without corresponding velar ones. In these languages the pronunciation of the stops may show some variation between the uvular and velar locations, but the uvular is more typical.

The most common type of uvular continuant is a voiceless fricative, for which the phonetic symbol is [χ], based on the shape of the Greek letter "chi". In some varieties of German the sound at the end of a word such as *Bach* 'stream' is a sound of this type (in other varieties it is a velar fricative). A total of 51 of the languages have voiceless uvular fricatives. The largest cluster of these is found in the north-west of North America, where there is an area in which almost all the languages have both uvular stops and uvular fricatives. Tlingit (Na-Dene), the major indigenous language of southeastern Alaska, has ten different uvular consonants. There are three kinds of stops—voiceless, aspirated, and ejective—and two kinds of fricatives—voiceless and ejective—and each of these types occurs with and without the rounding of the lips called labialization (written in phonetic transcription with a small raised "w" after the main symbol). Examples of words with each of these sounds are given in Table 1 below.

Table 1 Words exemplifying the ten uvular consonants of Tlingit (Maddieson et al. 2001)

voiceless plosive	*qák^{w}*	'tree spine'
aspirated plosive	*q^{h}ák^{w}*	'basket'
ejective stop	*q'ak^{w}*	'screech owl'
labialized voiceless plosive	*náaq^{w}*	'octopus'
labialized aspirated plosive	*q^{wh}áan*	'people, tribe'
labialized ejective stop	*q^{w}'átɬ*	'cooking pot'
voiceless fricative	*χaak^{w}*	'fingernail'
ejective fricative	*χ'áak^{w}*	'freshwater sockeye salmon'
labialized voiceless fricative	*χ^{w}astáa*	'canvas, denim'
labialized ejective fricative	*χ^{w}'áaɬ'*	'down (feathers)'

The other area of the world where languages with an especially large number of distinct uvular consonants are found is the Caucasus. For example, in Archi (Nakh-Daghestanian; eastern Caucasus), there are sixteen different types of uvular consonants. Both voiced and voiceless uvular fricatives occur in this language, and besides

uvulars with labialization, ones with an added narrowing in the pharynx are also found.

Both German and French have a more unusual type of voiced uvular continuant, represented by the letter *r* in the orthography, in words such as German *rot* 'red' and French *rue* 'street'. In the most careful speech this sound is a uvular trill, during which the uvula makes several rapid contacts with the back of the tongue. The trilling is created purely by establishing the right aerodynamic conditions, which will cause the uvula to vibrate as the air flows past its location at the back of the mouth. Since several factors have to be within critical limits for the vibration to occur it is not surprising that trilling quite often fails to take place and instead a uvular fricative or approximant is pronounced. For many speakers of standard French or German one of these variants has become the norm.

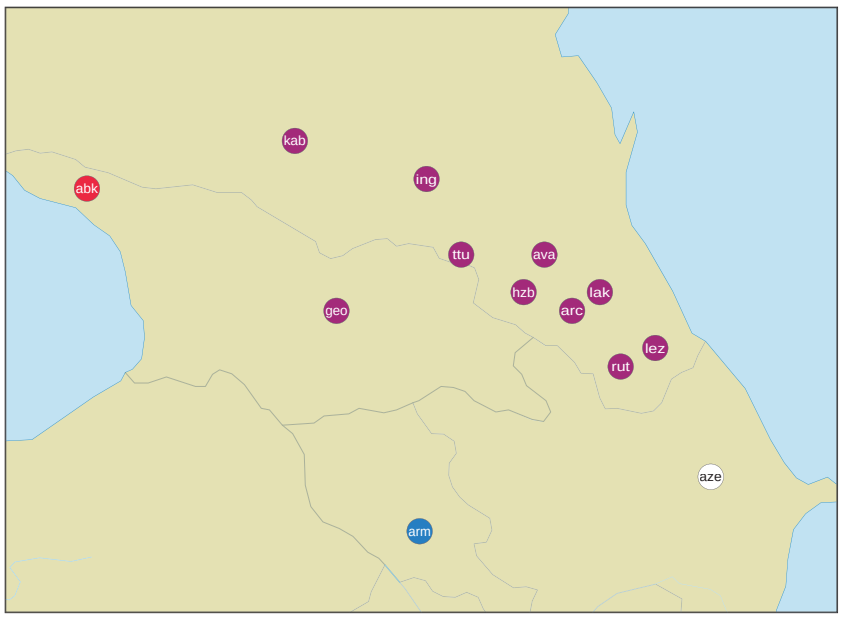
One further kind of uvular continuant is found in a few languages, a voiced uvular nasal (written in phonetic transcription with the small capital letter [ɴ]). In the present survey only Japanese is considered to have a uvular nasal in its inventory of consonants, but it is a special case and other interpretations are possible. This nasal is the only consonant which can appear in word-final position in Japanese, as in the word /hoɴ/ 'book'. Since this nasal has a clearly different distribution and pronunciation from the syllable-initial nasal sounds /m, n/, it is usually regarded as a distinct sound in discussions of the sound system of Japanese. Uvular nasals also occur in the dialects spoken by the Inuit peoples of Alaska, Canada, and Greenland. The facts are complex but, as described by Rischel (1974) for West Greenlandic, in most cases the uvular nasals seem to be the result of a nasal consonant and a uvular stop "fusing" together into a single sound. For this reason, the uvular nasal is not considered a basic sound of this language although uvular stops and fricatives are.

The strong geographic concentration of languages with uvulars in the north-west of North America connects across the Bering Strait to an area in the far north-east of Asia where a cluster of languages with uvular stops and continuants occurs. Apart from the languages in the Eskimo-Aleut family, which account for only two of the languages in these areas in the sample, the languages on the Asian and the American sides of the Strait have not yet been convincingly shown to be related to each other. However, since the Bering Strait was undoubtedly an important route for the entry of human beings to the New World, typological similarities across the strait are suggestive of ancient or continuing contacts between the continents.

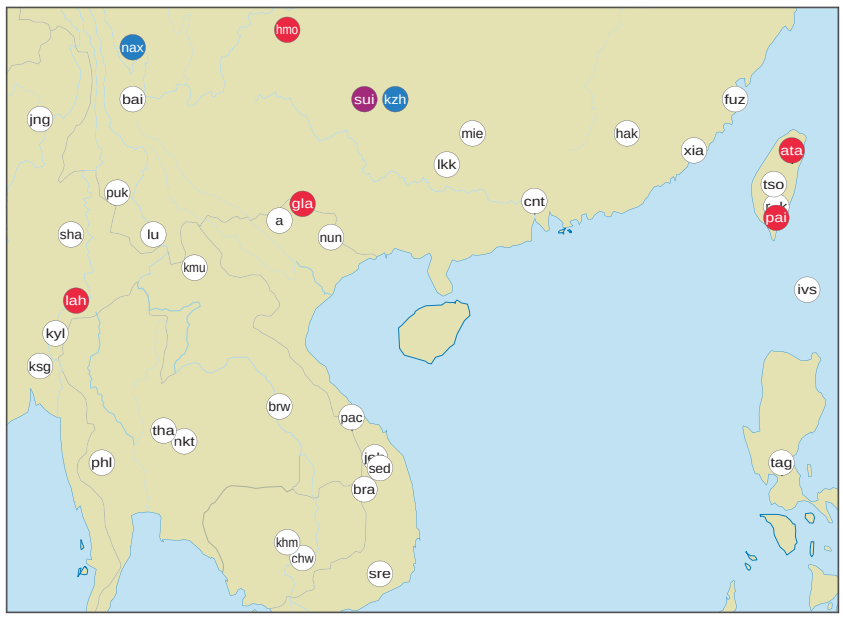
6 Uvular Consonants

IAN MADDIESON

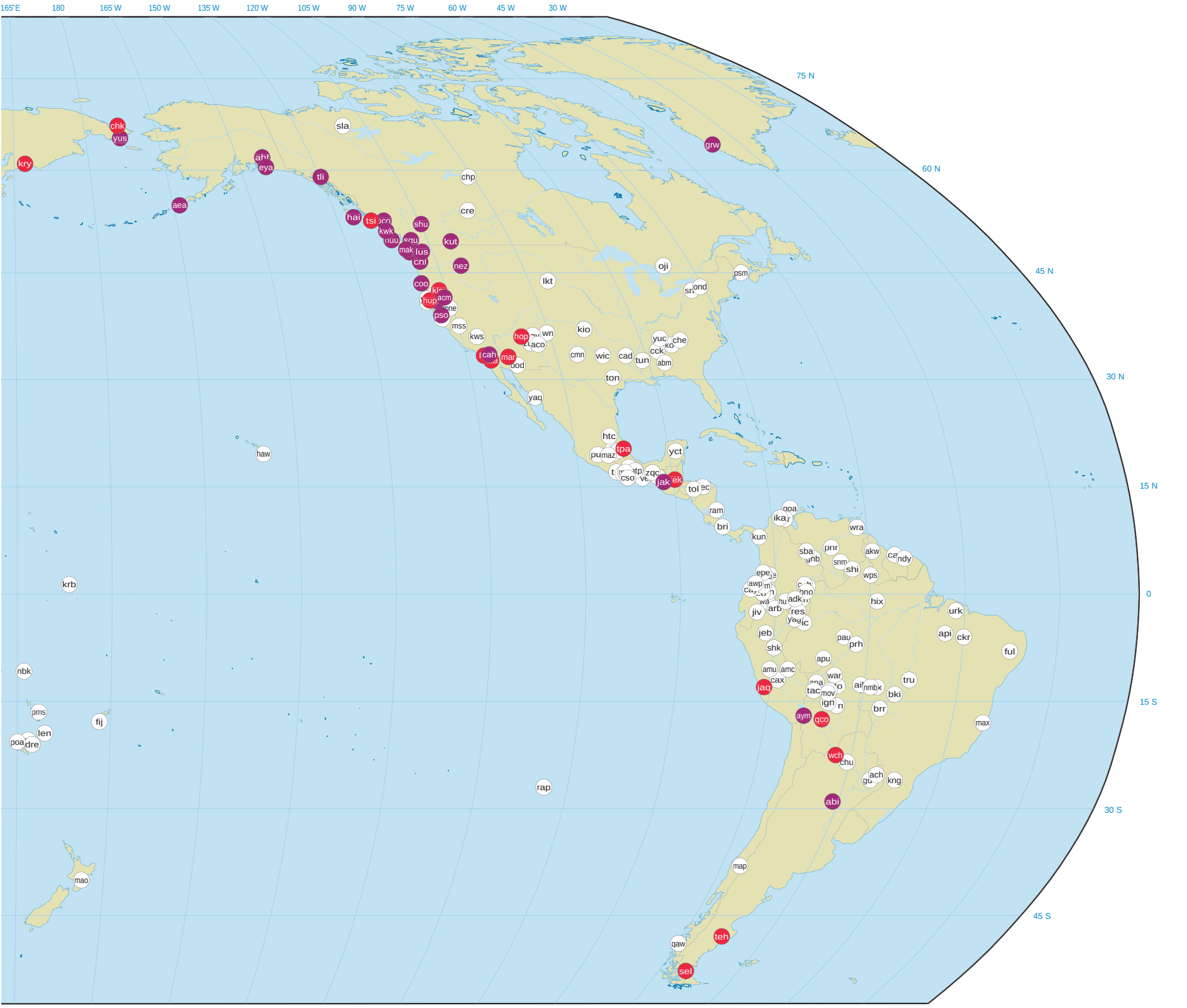
None [468]
Uvular stops only [38]
Uvular continuants only [12]
Uvular stops and continuants [48]

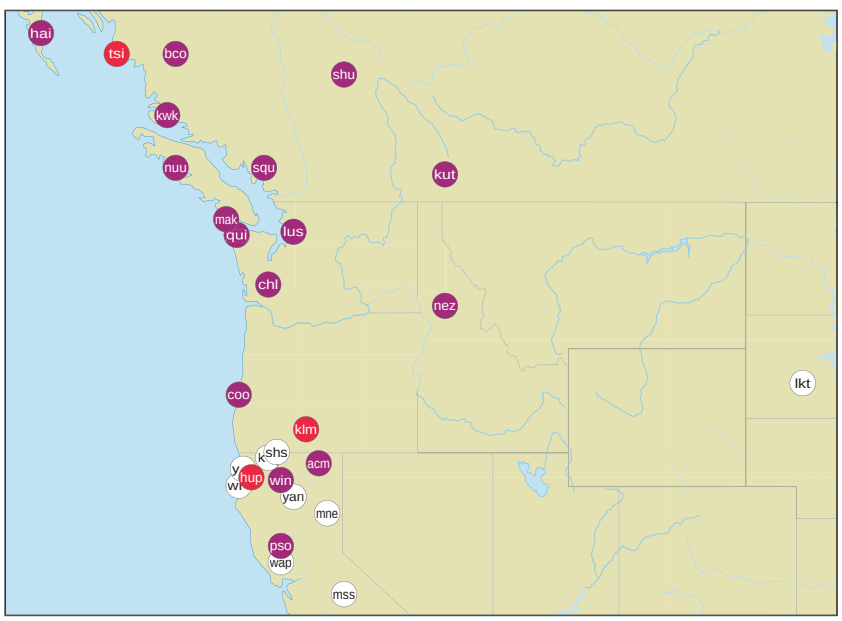


Caucasus

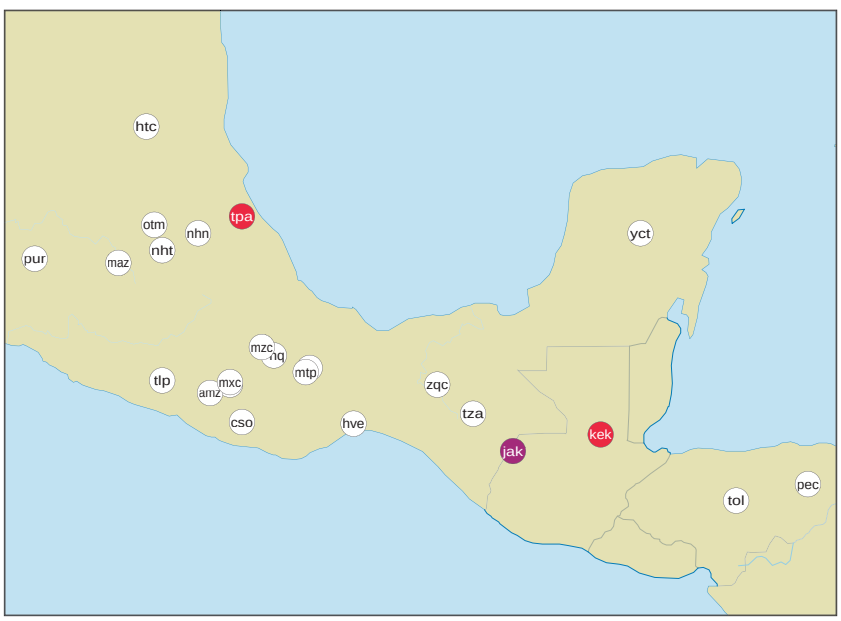


Indochina and surrounding areas





western North America



Mesoamerica

7 Glottalized Consonants

IAN MADDIESON

1 Introduction

In this chapter the distribution of consonants whose production involves certain special actions of the larynx will be discussed. In the great majority of consonant sounds, the vocal folds inside the larynx are either open so that air from the lungs can freely flow into the mouth, or they are brought close together so that air flowing between them causes them to vibrate, producing the effect called **voicing** (as discussed in Chapters 1, 4, and 5). All languages have some sounds which are voiceless, that is, are produced with open vocal folds, and some which are voiced. In addition to these two settings there are other, less commonly used adjustments of the larynx involving either a tighter constriction of the vocal folds and/or a movement of the larynx a short distance up or down in the throat. Consonants with these characteristics are known as **glottalized** (from the word *glottis*, the name given to the space between the vocal folds).

The glottalized consonants can be conveniently grouped into three classes. The first of these classes most typically involves a complete closure of the vocal folds followed by an upward movement of the larynx. If these movements are made at the same time as a closure in the mouth is maintained as for a stop, the air in the space between the closed vocal folds and the mouth closure will be compressed as the distance between the two closures is reduced. When the mouth closure is released a characteristically sharp explosive noise is created. Stops of this kind are known as **ejective stops**. They are written in phonetic notation with a raised apostrophe after the symbol that is employed to represent the mouth action; thus /pʼ/ represents a bilabial ejective stop, /kʼ/ a velar ejective stop, and so on. For example, Wintu (Penutian; California) has word-roots such as /pʼat-/ 'fishtail' and /kʼil-/ 'hail' contrasting with /pat-/ 'mountain lion' and /kil-/ 'liver'. The ejective mechanism is also often used in producing affricates, as in the Chulupí (Mataco-Guaicuru; Paraguay) word for 'milk' /tsʼoʔs/, and more rarely for producing fricatives. Tlingit (Na-Dene; Alaska) is among the relatively few languages with ejective fricatives, as in the word /sʼaaw/ 'crab'.

The second group of glottalized consonants are stops in which the critical movement of the larynx is downward rather than upward. If the downward movement is sufficiently vigorous, air will briefly flow into rather than out of the mouth when the mouth closure is released. Because of this possibility of inward airflow, this class of sounds is known as **implosives**. Most typically the vocal folds are in the voicing position in the production of sounds of this group, but they may be more tightly closed. The phonetic symbols for these sounds are the letters used for voiced stops modified by the addition of a hook to the right at the top, as in /ɓ, ɗ, ɠ/. For example, in Murle (Surmic, Nilo-Saharan; southern Sudan) the word for 'crocodile' is /aɠul/. Similar sounds are often referred to as "preglottalized voiced stops" by linguists working on Asian and Pacific languages.

The third group of glottalized consonants are **glottalized resonants**. Resonant (also known as sonorant) consonants are those produced with an unimpeded flow of air through the mouth or the nose, such as those at the beginning of the English words *lot*, *not*, *rot*, *yacht*. Consonants of this general type are usually voiced in the normal manner, but they can be produced with a closer constriction of the vocal folds which interrupts or modifies the normal voicing. Glottalized resonants are often symbolized with a following apostrophe in the same way as ejectives. For example, in Yapese (Austronesian; Micronesia) the word for 'sick' begins with a glottalized bilabial nasal and can be written /mʼaar/. However, glottalized resonants are not produced with either a raising movement of the larynx or a lowering movement of the larynx.

As with any grouping into categories, some difficult decisions have to be made about which cases to include or exclude. For example, stops which are described as having a closure of the glottis accompanying the closure in the mouth but which lack the larynx-raising characteristic of an ejective stop have been grouped together with ejectives. For this reason Yurok (Algic; California) is among the languages included in the survey as having ejectives, even though the glottalized stops in this language do not seem to require larynx raising. Korean is also included because it has a set of consonants that seem to be produced with a closer than usual vocal-fold position, although they are not otherwise reminiscent of ejectives. This series of consonants, written with doubled letters in the Korean orthography, is often called "fortis" in the literature on Korean. It should thus be borne in mind that the terms **ejective** and **implosive** are being used here to refer to somewhat more inclusive classes of consonants than is traditional in the phonetic literature.

In order to show the co-occurrence patterns between the classes of glottalized consonants, a quite complex categorization scheme for languages with these consonants has been adopted. Languages with ejectives only, with implosives only, and with glottalized resonants only each form a separate category. Those whose consonant inventories include members of more than one of these classes of consonants are divided into those with only ejectives and implosives, those with only ejectives and glottalized resonants, those with only implosives and glottalized resonants, and those with all three classes of glottalized consonants. Languages with no glottalized consonants of any sort form the first class of languages shown on the map.

○	1. No glottalized consonants	412
●	2. Ejectives only	57
●	3. Implosives only	55
●	4. Glottalized resonants only	3
●	5. Ejectives and implosives	13
◆	6. Ejectives and glottalized resonants	19
◆	7. Implosives and glottalized resonants	4
◆	8. Ejectives, implosives, and glottalized resonants	3
	total	566

Two types of consonants which might have been included are not counted among the glottalized consonants. The first is the consonant known as the **glottal stop**, for which the phonetic symbol is /ʔ/. This is the result of a brief closure of the vocal folds without any accompanying movement of the tongue, lips, or other speech organs. This sound is familiar as a marker of the colloquial English of London, where it occurs in place of the sound /t/ between a stressed and an unstressed syllable in words such as *letter* and *little*. In other languages a glottal stop is used to keep certain adjacent vowels clearly separate from each other, as in German *geübt* /gəʔypt/ 'practised'. The glottal stop occurs in many languages—in fact in a much larger number of languages than the glottalized consonants which are the focus of this chapter—as a full member of the set of consonants. For example, in Hausa, the main language of northern Nigeria and much of the eastern Sahel region of Africa, /saaʔàa/ 'hour' is a quite different word from /saatàa/ 'theft'. Glottal stops may also occur with what are called "secondary articulations" such as labialization, so that /ʔʷ/ has a similar modification to that found in /kʷ/; similarly for palatalized /ʔʲ/. The sound /ʔʷ/ is obviously quite similar to the glottalized resonant /wʼ/, and either of these sounds might be analysed as a sequence of two consonants, /ʔ/ and /w/. Arguments must be sought to determine which of the potential interpretations yields a more satisfactory overall view of

the language's structure. Because it has a contrastive glottal stop and other palatalized consonants, Hausa is here interpreted as having a palatalized glottal stop rather than a glottalized palatal resonant; thus the word for 'daughter', orthographic *'ya*, is interpreted as /ʔʲaa/ rather than /jˈaa/.

The second excluded class is what are known as **breathy voiced consonants**; these have a special position of the vocal folds in which they are a little wider apart than the position for normal voicing. Hindi is the best-known language with such sounds. Neither glottal stops nor breathy voiced consonants have traditionally been viewed as belonging to the class of glottalized consonants.

2 Co-occurrence in inventories

At least some glottalized consonants occur in the consonant inventories of 154 of the 566 languages surveyed for this chapter, that is, in a little over a quarter of the languages (27.2 per cent). Among the three classes of these consonants as defined above, ejectives are more widely found than implosives, and glottalized resonants are the least widespread. Ejectives or ejective-like consonants occur in 92 (16.3 per cent) languages in the survey, implosives or implosive-like consonants occur in 75 (13.3 per cent), and glottalized resonants in just 29 (5.1 per cent). Table 1 highlights the patterns of occurrence and co-occurrence among the three classes of glottalized consonants in languages having sounds in one or more of these classes, omitting only the handful which have only glottalized resonants.

Table 1 Co-occurrences of glottalized consonants of the three classes (three languages with glottalized resonants only not included)

	no Resonants	with Resonants
Ejectives, no Implosives	57	19
Implosives, no Ejectives	55	4
Ejectives and Implosives	13	3

These figures show that there is a very strong association between the occurrence of the ejective and the glottalized resonant classes in a consonant inventory, and a much weaker association between either of these classes and implosives. An overwhelming majority (135 out of 151) of the languages in the survey with implosives or with ejectives do not include members of the other class in their consonant inventories; on the other hand, a majority (22 out of 29) of the languages with glottalized resonants also have ejectives.

3 Geographical distribution

The geographical distribution of glottalized consonants is strongly regional. They are absent from most of western Eurasia, rarely found in northern Asia, the South Asian subcontinent, island Asia, and New Guinea, and are entirely unknown in Australian languages.

Ejectives are found most especially in the Americas, as well as in the more easterly and southern regions of Africa and in the Caucasus. Well-known languages of the Americas with ejectives include Cochabamba Quechua (Quechuan; Bolivia), Yucatec (Mayan; Mexico), and Navajo (Athapaskan; Arizona and New Mexico). They are particularly found along the Andean cordillera in the south, in Mexico and Guatemala, and in the more northwesterly parts of North America. Most strikingly, the consonant inventories of almost all the diverse indigenous languages of northern California, Oregon and Washington, British Columbia, the Yukon, and Alaska include ejectives. The occurrence of ejective-like consonants in Soqotri (Semitic; Yemen) is associated with the African area in which these consonants occur, which includes other Semitic languages in nearby Ethiopia and Eritrea such as Amharic and Tigre. The dense cluster of languages in the Caucasus with ejectives includes languages of four different families, Kartvelian (e.g. Georgian), North-West Caucasian (e.g. Kabardian), Nakh-Daghestanian (e.g. Ingush, Lak, Archi), and Indo-European in the form of Eastern Armenian. The three languages with ejectives shown outside these areas are Korean, Itelmen (Chukotko-Kamchatkan; Siberia), and Yapese. The Korean consonants in question are not true ejectives, but those in Itelmen and Yapese are, showing that consonants of this type can occasionally develop in geographical isolation.

Languages with implosives are primarily found in Africa outside the northernmost tier, and in South-East Asia. A few cases occur in the Americas and elsewhere. The African languages with implosives as their only class of glottalized consonants are found across the "middle belt" from the west to the east of the continent, and include languages of three main families, Niger-Congo (e.g. Kpelle, Gwari, Kisi), Nilo-Saharan (e.g. Maasai, Murle, Yulu), and the Chadic branch of Afro-Asiatic (e.g. Ngizim, Kera). The major area within which languages having both ejectives and implosives are found is eastern and southern Africa, and a high proportion (11 out of 16) of the languages in which the two types co-occur come from this broad region. These languages are drawn from all four of the major language families of Africa, namely Niger-Congo (e.g. Zulu), Afro-Asiatic (e.g. Kotoko, Hamer, Dahalo), Nilo-Saharan (e.g. Ik, Komo), and Khoisan (e.g. Deti). The South-East Asian area within which implosive-like consonants occur includes members of the Tai-Kadai (e.g. Sui, Lakkia) and Austro-Asiatic (e.g. Pacoh, Vietnamese) families. Implosives, alone and in co-occurrence with ejectives, therefore display a primarily areal rather than genealogical pattern of distribution.

Glottalized resonants occur most often in the languages of the Americas. Of the 29 languages in the survey with consonants of this class, 20 are found in the Americas, and of these, 16 are found in North America north of 34° N and west of 75° W. Several distinct families, including Salishan (Squamish, Lushootseed), Wakashan (Kwakw'ala, Nuuchahnulth), Keresan (Acoma), Algic (Yurok), and Na-Dene (Slave), are represented. In the overall sample, glottalized resonants are found in only two of the African-area languages with ejectives (!Xóõ and Soqotri), and in none of those in the Caucasus. The areal restriction suggests that the association between glottalized resonants and ejectives might best be viewed as a result of overlapping patterns of spread in a single area, and not as the consequence of any particular linguistic dependence between the occurrence of these two classes of consonants.

4 Discussion

Because their production involves a more intricate coordination of the actions of the larynx with the actions of the articulators in the mouth than many of the more common sounds, glottalized consonants are usually regarded as being inherently complex sounds. As noted in Chapter 1, a general principle has been proposed according to which more complex consonants are predicted to be more likely to occur in languages with larger numbers of contrasting consonants (the size principle). We can see that this prediction is generally supported by comparing the maps in this chapter and in Chapter 1. Table 2 summarizes the distribution of languages with glottalized consonants across the five classes of consonant inventory size established in that chapter.

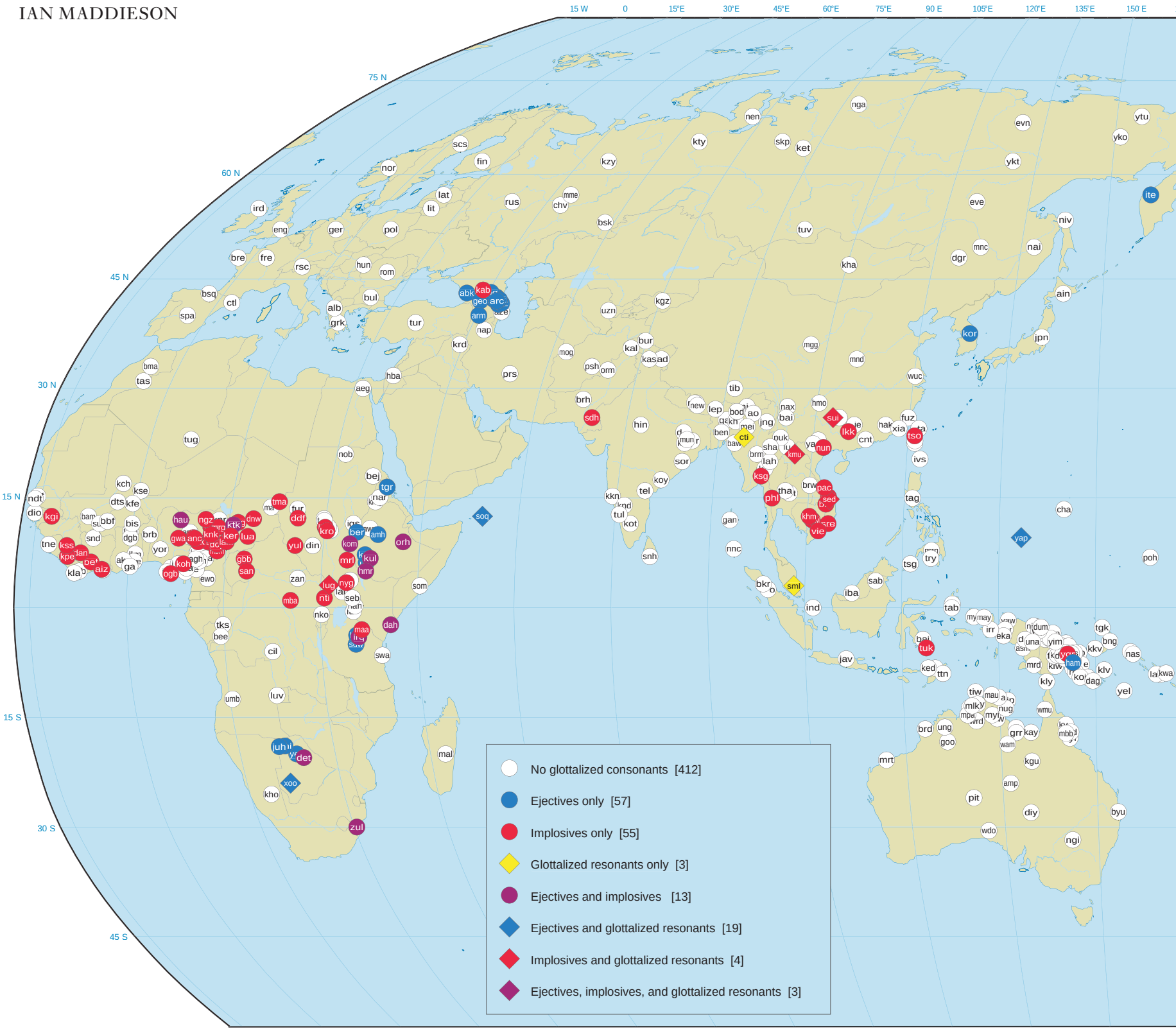
The table shows the percentage of the languages falling within each consonant inventory size class which have one or more glottalized consonants among the consonants constituting their inventory. Only a few of the languages with small consonant inventories have any glottalized consonants, whereas two-thirds of those with large inventories include one or more glottalized consonants, and the proportion increases with each increase in overall inventory size. Thus the predictions of the size principle are strongly supported by these data, as well by related data to be presented in Chapter 19.

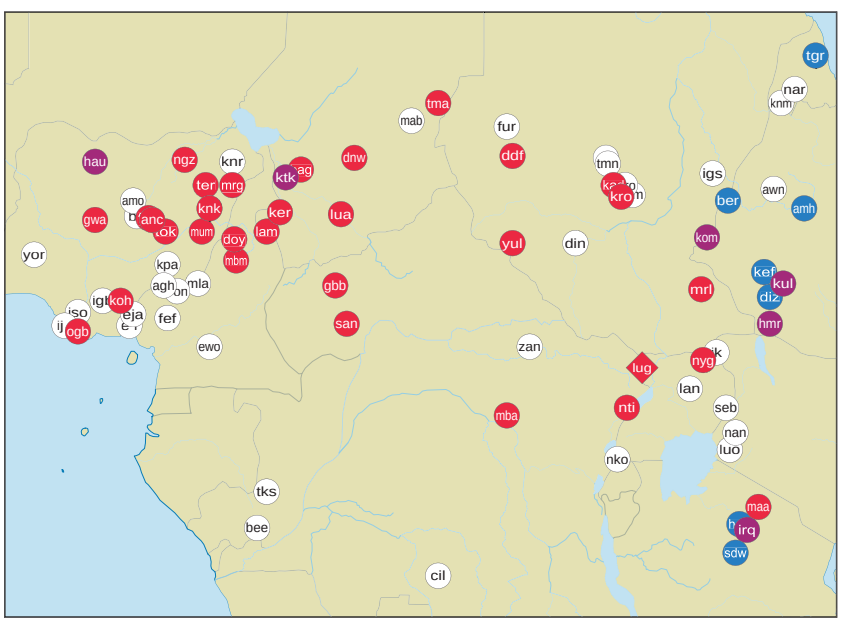
Table 2 Languages with glottalized consonants, by consonant inventory size

Consonant inventory size class	per cent with glottalized C's
small	8.7%
moderately small	10.7%
average	21.5%
moderately large	39.3%
large	66.7%

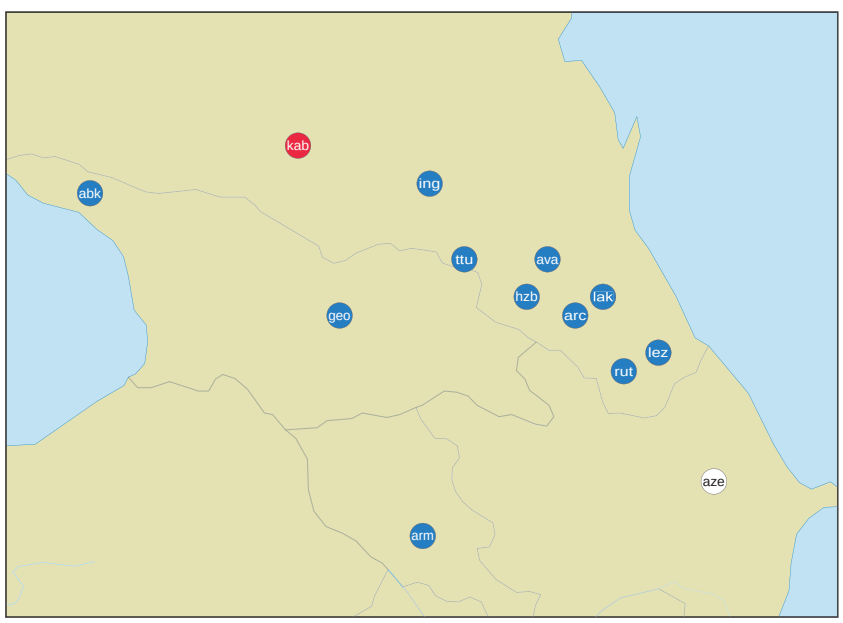
7 Glottalized Consonants

IAN MADDIESON

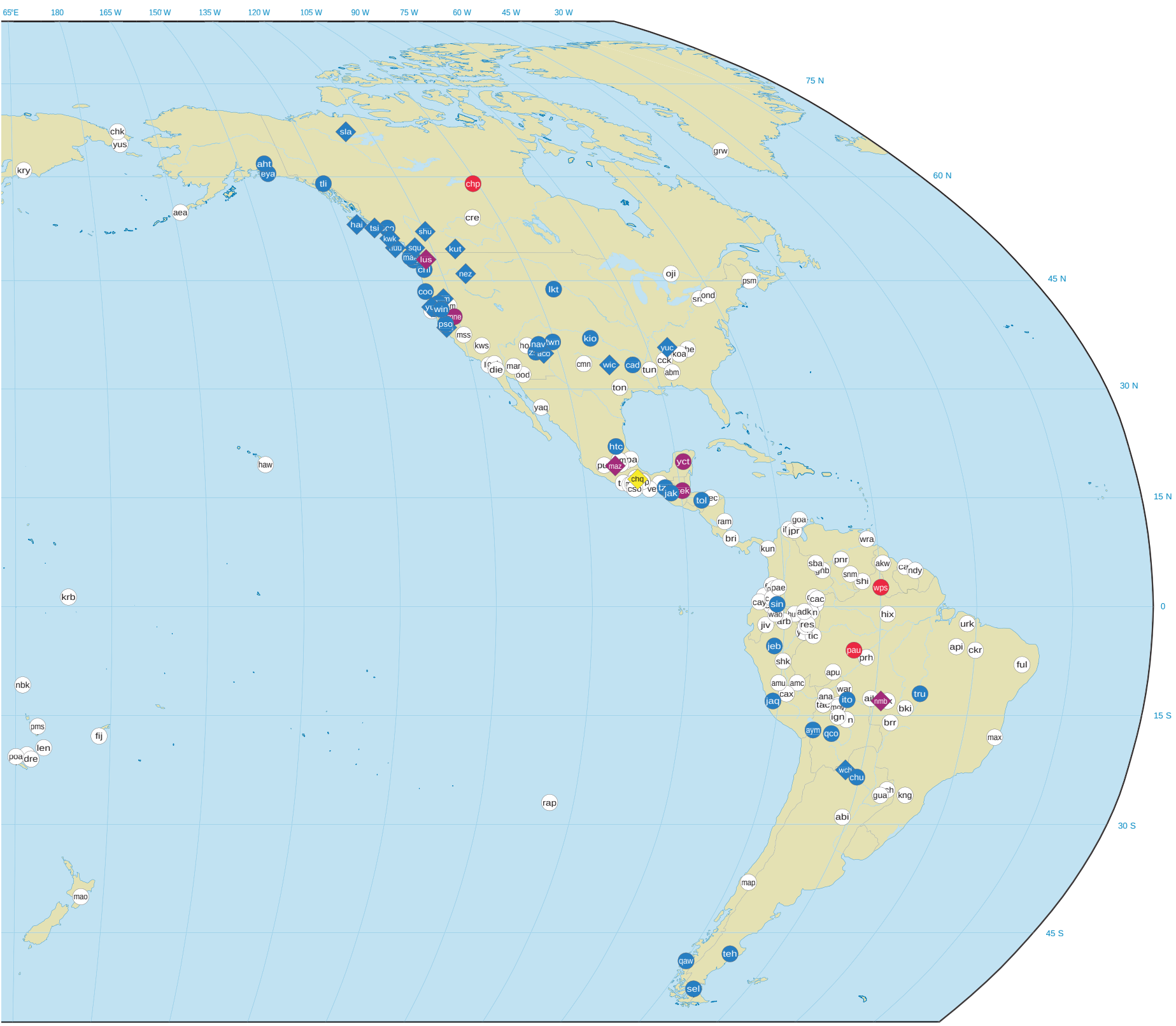


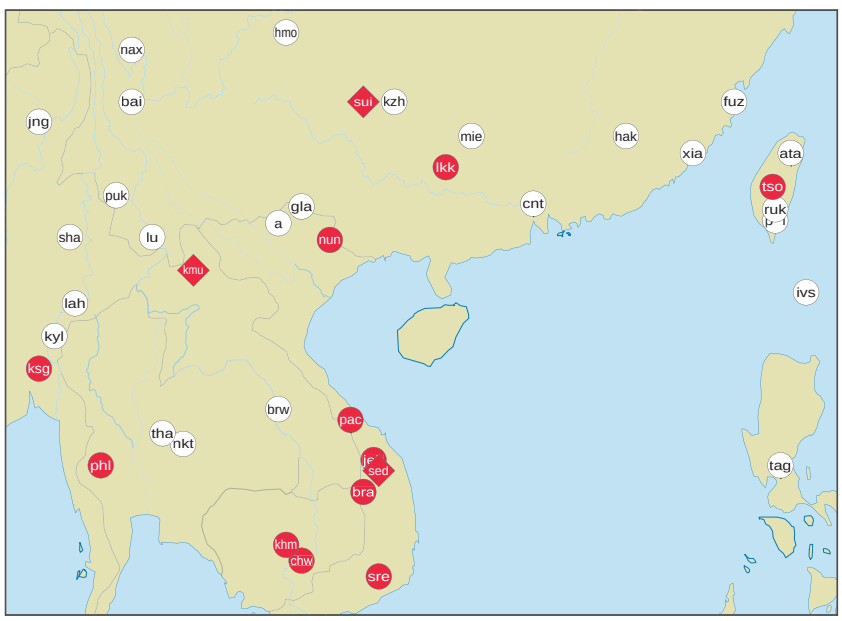


central Africa

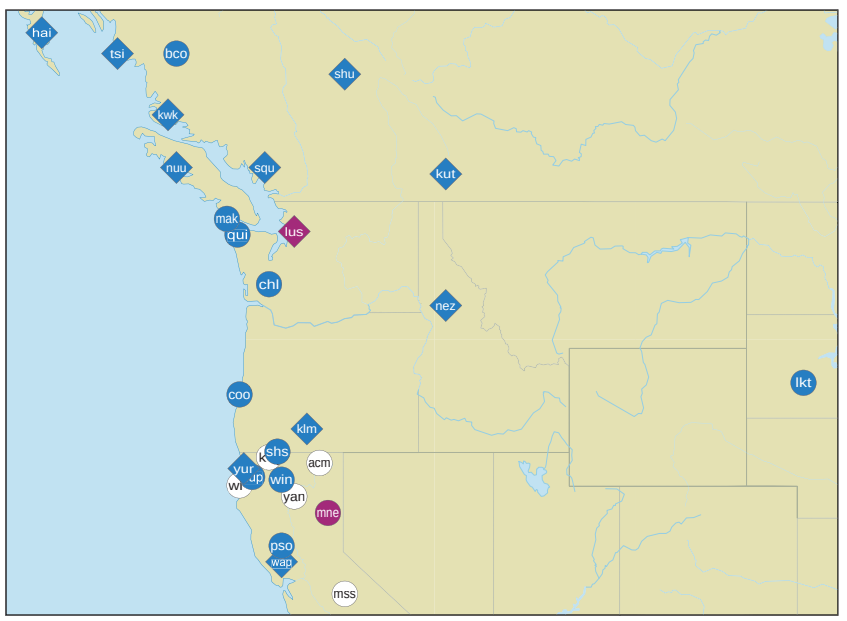


Caucasus





Indochina and surrounding areas



western North America

8 Lateral Consonants

IAN MADDIESON

1 Introduction

In their consonant inventories most of the languages in the world have a consonant sound similar to that represented by the *l* in English *lake*, Spanish *lobo* 'wolf', or Indonesian *laki* 'husband'. The technical phonetic description of sounds of this type is that they are voiced lateral approximants produced at the dental or alveolar place of articulation. The word *lateral* in this description refers to the fact that during their production the flow of air originating from the lungs is directed around a region of contact between the tongue and the roof of the mouth which leaves a passage for the air to escape to one or both sides of this central closure. The size of the channel through which the air flows is wide enough that the flow is unimpeded and does not create friction noise as it passes, and the vocal folds are in the normal position to produce voicing vibrations in the same way as is typical during vowel sounds. Sounds with lateral airflow are most often produced with the tip or blade of the tongue making contact with the back of the upper teeth or at a position just behind the teeth; hence they are labelled dental or alveolar. If you pronounce the "l-sound" of a language familiar to you but then keep the tongue in the same position and draw the breath quickly inwards the colder inward-flowing air will help you feel where the side passage for airflow is located.

In this chapter the occurrence of laterals of this common type and of various other kinds of lateral consonants that are less frequently found will be described, as well as the distribution of languages that have no lateral consonants. As in other chapters discussing aspects of consonant inventories, the set of sounds under consideration is the set of contrasting sounds as represented by their most typical pronunciation (see Chapter 1).

2 Classification and distribution of laterals

Over four-fifths (471 or 83.2 per cent) of the languages surveyed for the chapter have one or more lateral consonants, while the remainder (95 or 16.8 per cent) have no lateral segments of any kind in their consonant inventory. These languages make up the "no laterals" class. The largest number of these occurs in the northern half of South America, with a significant cluster also occurring in New Guinea, and a scattering of other cases in other parts of the world. About a third of these languages with no laterals have no "liquid" consonants of any kind. Liquids are the class of consonants including both lateral approximants and the family of sounds usually represented by the letter *r*. The latter are sometimes called "rhotics", after the name of the Greek letter "rho" which is used to write a member of this family. Laterals and rhotics form a class together because in many languages these sounds have a special freedom to occur in consonant clusters (e.g. in English *play*, *pray*; *clue*, *crew*). Also, in some languages laterals and r-sounds are alternative ways of pronouncing the same distinctive unit in the sound system, e.g. in Korean. Among languages with no liquids are Apurinã (Arawakan; Brazil) and Siona (Tucanoan; Ecuador); the New Guinea languages Dera (Senagi), Usan (Madang, Trans-New Guinea), and Gadsup (Chimbu, Trans-New Guinea); Hopi (Uto-Aztecan; Arizona), Seneca (Iroquoian; New York State), and Tlingit (Na-Dene; Alaska) among the native languages of the United States; Efik (Cross River, Niger-Congo; Nigeria) and Xiamen (Chinese; southeastern China).

○	1. No laterals	95
●	2. /l/, no obstruent laterals	388
●	3. Laterals, but no /l/, no obstruent laterals	29
●	4. /l/ and lateral obstruents	46
●	5. No /l/, but lateral obstruents	8
	total	566

Languages with lateral consonants are divided here first according to whether they have a lateral of the typical type, that is, a dental or alveolar voiced lateral approximant as normally represented by the letter *l*. Secondly they are divided according to whether they have lateral fricatives or affricates. Like other fricatives, lateral fricatives are sounds in which the channel through which the air flows is narrowed to the point that the flow of air becomes turbulent and noisy. However, in this case the narrowed channel is to one side or the other of a contact between the tongue and the teeth or the roof of the mouth. Welsh, a language that does not figure in the sample, is well known as an example of a language with a lateral fricative, which is represented by *ll* in the orthography in words such as *llan* 'church'. The phonetic symbol for this sound, a voiceless alveolar lateral fricative, is /ɬ/; thus the word can be transcribed /ɬan/. Lateral affricates are stops in which a /t/ or /d/-like closure is released into a lateral fricative by lowering one side of the tongue. As discussed in Chapter 4, fricatives and affricates are included in the larger class of sounds called obstruents, so the two types of lateral sounds in question can be jointly referred to as lateral obstruents.

Just over three-quarters of the languages surveyed (434 or 76.7 per cent) have a lateral consonant of the typical type. The great majority of these, 388 or 68.6 per cent of the total sample, do not have any obstruent laterals. These languages are shown on the map as the "/l/, no obstruent laterals" class. This pattern is dominant throughout Africa, Australia, Europe, and Asia with the exception of a few areas such as the Caucasus region and small parts of Africa. As the frequency of languages of this type indicates, it is most typical for the consonant inventories of the world's languages to include /l/ and to exclude obstruent laterals. Some of the languages in this class also have other lateral sounds which are not obstruents, such as ones made with a constriction further back in the mouth, for example at the palatal place of articulation, or pronounced without voicing or with glottalized voicing (see Chapter 7). The sound represented by *ll* in Spanish orthography in words such as *calle* 'street' is pronounced as a voiced palatal lateral approximant (phonetic symbol /ʎ/) in some varieties of Spanish, including the variety of Peninsular Spanish included in the sample. A voiceless alveolar lateral approximant occurs in Standard Spoken Tibetan in the word for 'god' /l̥á/, which forms the first part of the place-name Lhasa.

In addition to the languages of the most common type "/l/, no obstruent", there are a further 29 languages in the sample which have one or more nonobstruent lateral consonants and have no obstruent laterals, but which lack a lateral of the typical type. The most common situation in these languages is that they have one lateral consonant which is pronounced as a flap, a sound with a much shorter duration than the typical lateral approximant. There are, however, a range of other possibilities, such as having one or more lateral sounds made further back in the mouth, or with other

than normal voicing. The languages in this rather mixed group are named the "no /l/, no obstruent laterals" group.

Obstruent laterals occur in 54 of the languages surveyed (9.5 per cent). These languages have been divided according to whether they do or do not also have /l/. Those with /l/ are the "/l/ and lateral obstruents" group. Those which lack /l/ are the "no /l/, but lateral obstruents" group. Languages with lateral obstruents are geographically concentrated, occurring especially in the more northerly half of North America, particularly toward the west. Smaller clusters are found in the Caucasus, and around the Lake Chad basin and the north-east highlands regions of Africa. Lateral fricatives occur in 51 of these 54 languages. These sounds are most often voiceless—a voiced lateral fricative occurs in only 8 of the languages, and in every case but Tigak (Austronesian; New Ireland, Papua New Guinea) there is also a voiceless counterpart in the language. Three languages, Kabardian (North-West Caucasian; Russia), Yuchi (isolate; Tennessee), and Tlingit (Na-Dene; Alaska), have ejective lateral fricatives. Lateral affricates are less common than lateral fricatives. There are only 25 languages in the sample (4.4 per cent) with at least one lateral affricate; in 20 of these there is an ejective lateral affricate, making lateral affricates the class of consonants which are most likely to be produced with the ejective mechanism (see Chapter 7). In general, languages with any lateral affricates also have at least one lateral fricative in their consonant inventory; this is so for 22 of the 25 languages in question, with only Kiowa, Wintu, and Squamish, three languages of North America, reported as having affricates but no fricatives among their lateral sounds.

The great majority of the languages with obstruent laterals, 46 of the 54, also have a typical lateral approximant in their consonant inventory, leaving a small and interesting set of eight languages in which the common type of lateral is missing even though some have several obstruent laterals. The eight languages include four in the Pacific North-West, namely, Tlingit and Ahtna of the Na-Dene family, Nuuchahnulth of the Wakashan family, and Kutenai. Tlingit has five different types of obstruent lateral consonants, and Ahtna has four. These languages seem to have accentuated an areal pattern that favours the inclusion of obstruent laterals by excluding the common lateral type altogether from their inventories. A fifth member of this small class is also a North American language, Kiowa (Kiowa-Tanoan; Kansas and Oklahoma).

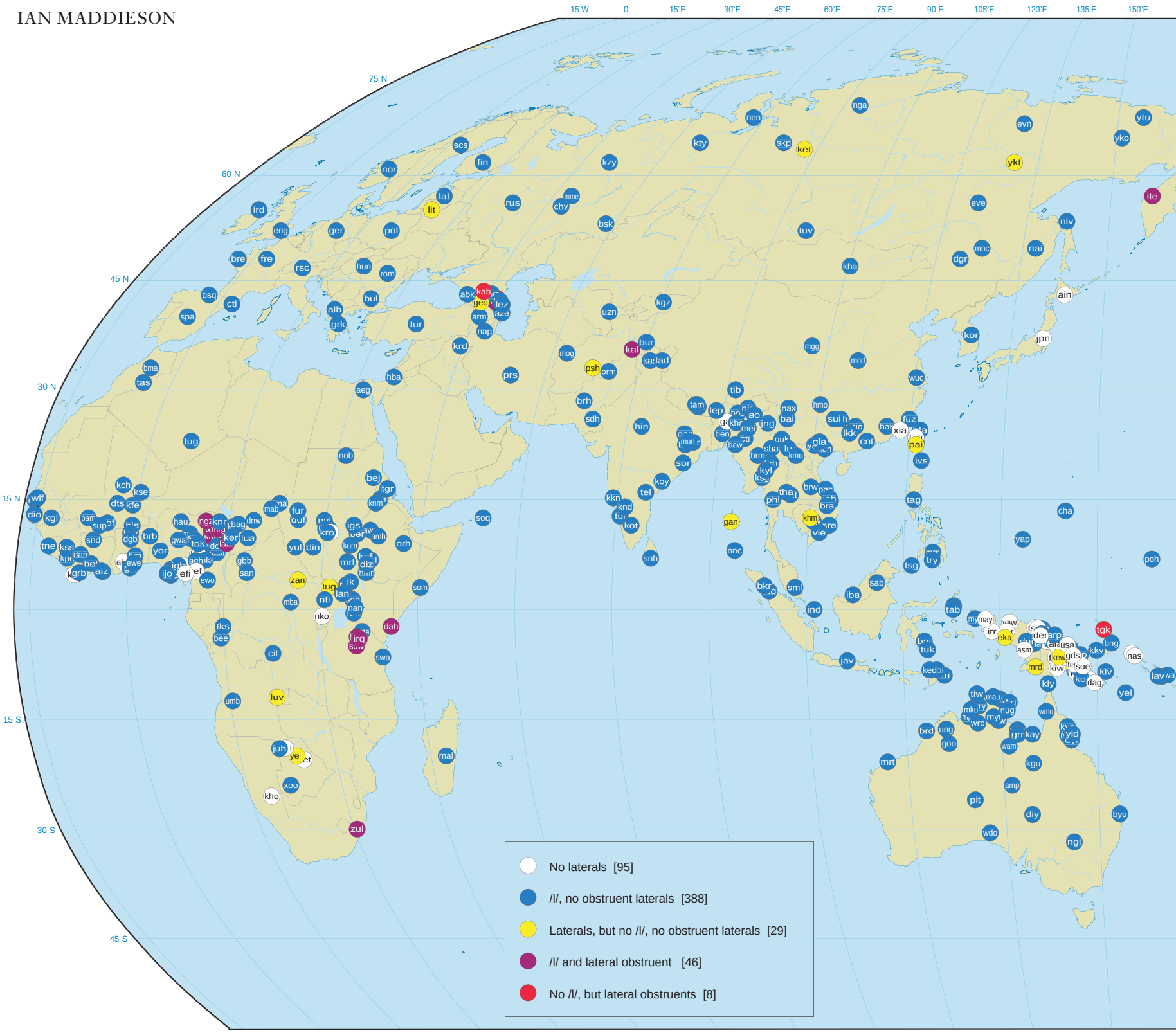
3 Discussion

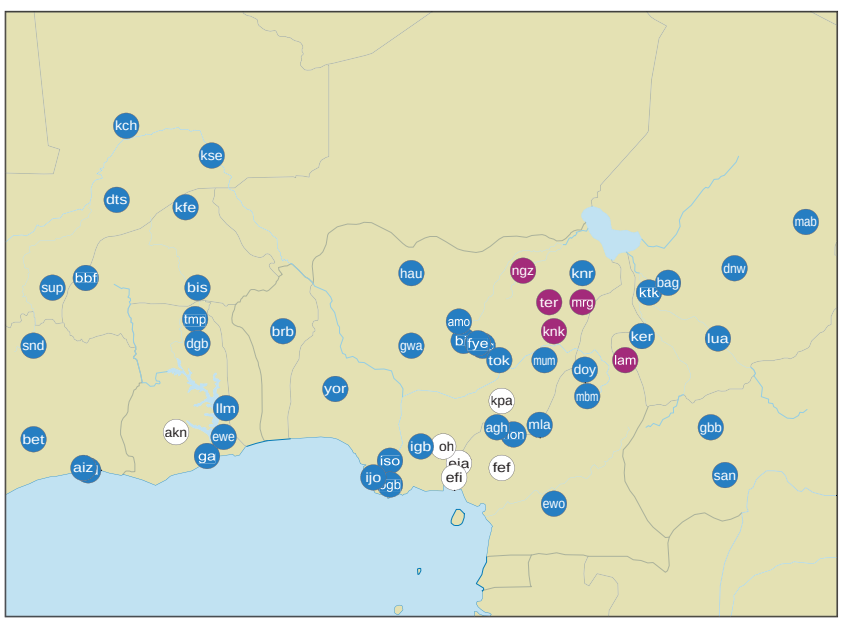
In summary, it appears at first glance that there is a very strong relationship among the main classes of laterals discussed in this chapter predicting how they will occur together in consonant inventories. If a language has a lateral affricate in its consonant inventory, then this generally entails the presence of a lateral fricative, and if a language has a lateral fricative, then this generally entails the presence of a lateral of the common approximant type. The first part of this observation makes a meaningful prediction, since 88 per cent of the languages with affricates also have fricatives; but the second part has no similar power to predict as the distribution is not greatly above a chance co-occurrence. Since /l/-type sounds are so commonly found, in a random subset of 51 languages drawn from our sample, 39 could be expected to have a sound of this type. Of the 51 languages in our survey which happen to have lateral fricatives, 44 also have a "typical" lateral approximant, a number not greatly above the chance level of 39. The occurrence of obstruent laterals therefore does not seem to be linked in any significant way to the occurrence of lateral approximants, but the two classes of obstruent laterals do show a strong tendency to occur together.

One final class of consonants with a lateral component should be briefly noted in this chapter. Among the click sounds discussed in Chapter 19 is a type of click in which the release of the closure is made by lowering a side of the tongue in much the same way as in a lateral affricate. Lateral clicks behave, in the phonology of the languages which have them, in very much the same way as other clicks and do not seem to have any connection to the consonants normally considered to represent the class of lateral sounds. They have therefore not been included among the sounds discussed in the present chapter.

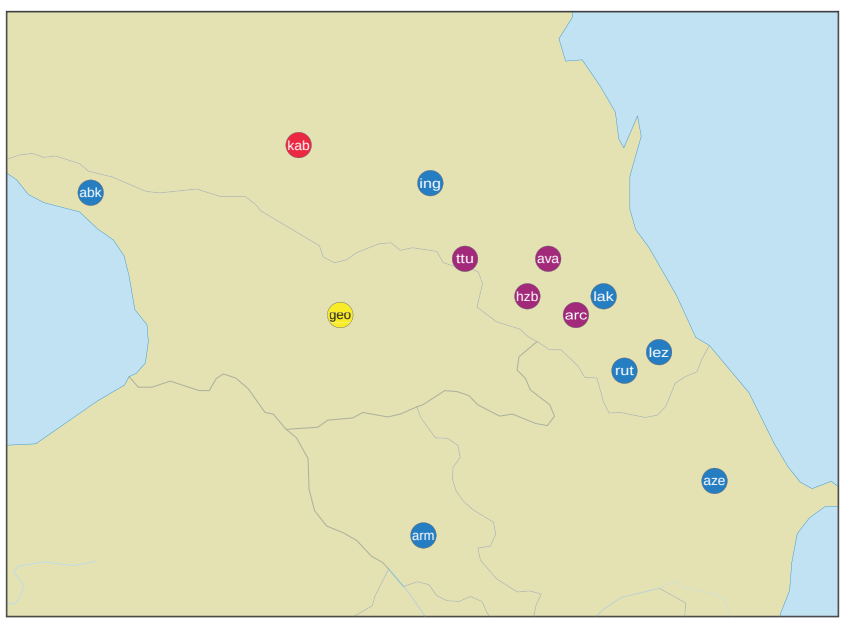
8 Lateral Consonants

IAN MADDIESON

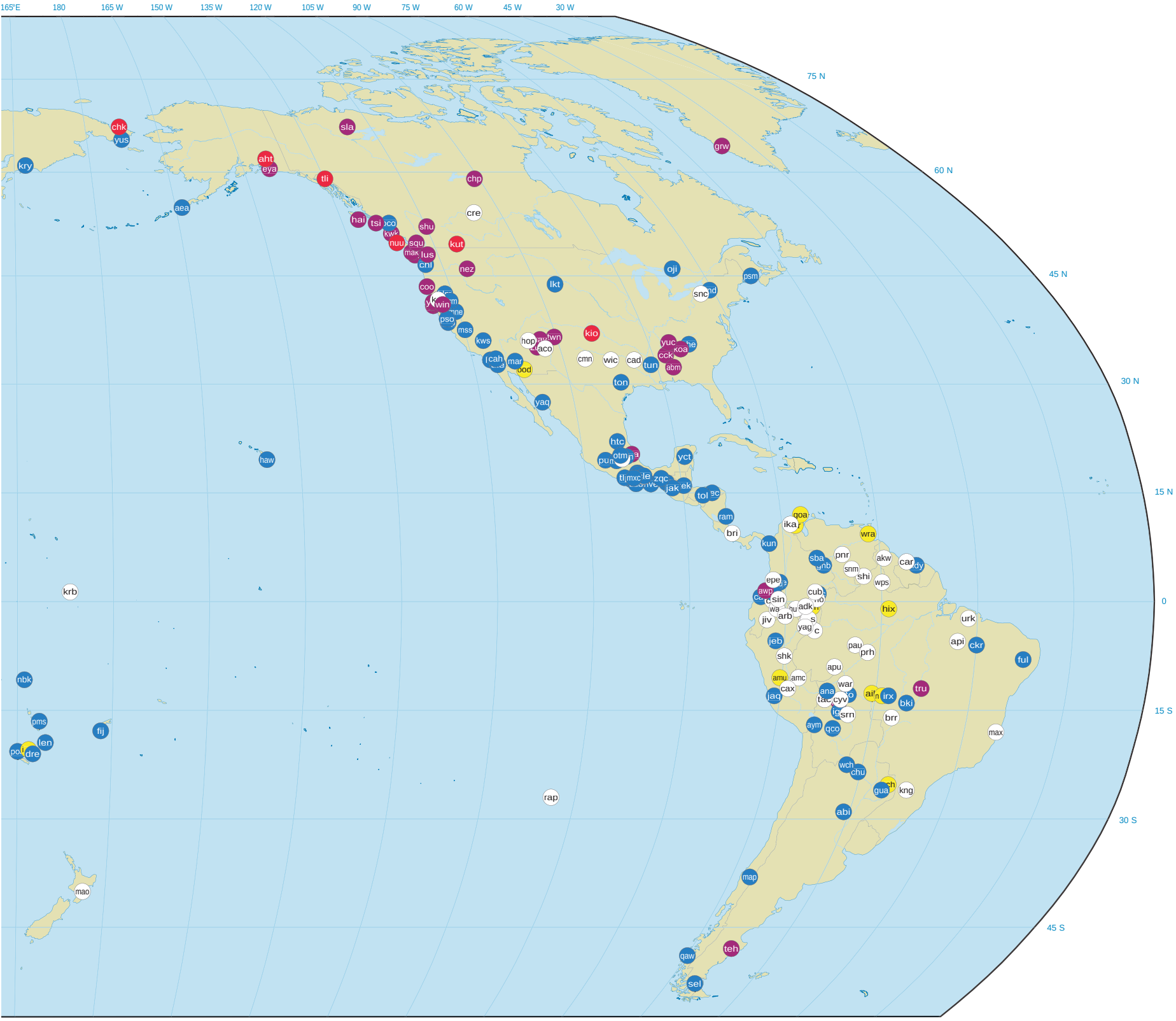


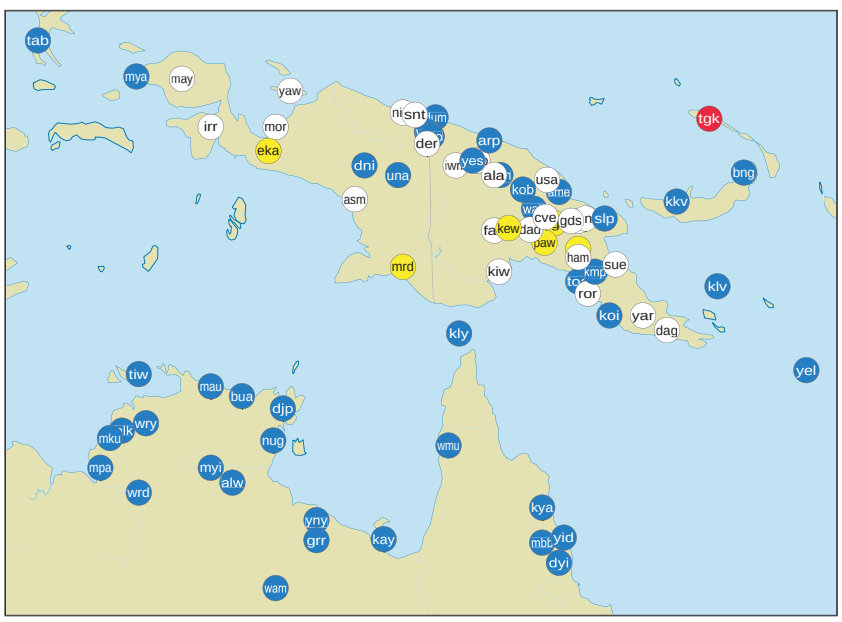


Nigeria and surrounding areas

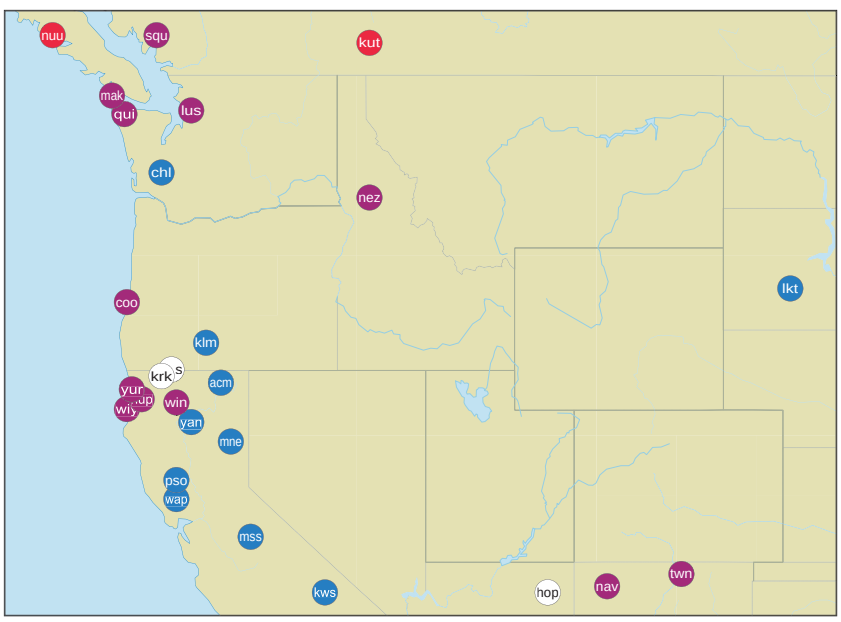


Caucasus





New Guinea and northern Australia



western United States

9 The Velar Nasal (*ŋ*)

GREGORY D. S. ANDERSON

1 Defining the values

This chapter addresses the distribution of the sound *ŋ* (the velar nasal) in the languages of the world. A very large number of languages make use of phonemic *ŋ*, while many others lack this sound. However, few of the sounds commonly found among the phonemic inventories of the world's languages exhibit a more clearly definable distribution than that exhibited by *ŋ*. This distribution has two unrelated aspects. One is the striking areal distribution of the presence versus absence of phonemic *ŋ* among the languages of the world. The other striking aspect of phonemic *ŋ* is its phonotactic distribution: in many languages possessing this sound, it may not appear in all positions in the word, but rather is restricted to initial, medial, or final position, or some combination thereof. In the case of restriction of *ŋ* to non-initial position, this, too, has a relatively pronounced areal skewing among the world's languages.

1. Velar nasal, also initially	146
2. Velar nasal, but not initially	88
3. No velar nasal	234
total	468

Assigning a given language to the group that possesses contrastive *ŋ* or the group that does not is not as straightforward as identifying *ŋ* within the sound inventory of the language—i.e. *ŋ* is often phonetically present but phonologically predictable or noncontrastive. Various issues complicate the determination of whether or not *ŋ* is phonemic (contrastive) in a given language. For example, the sound occurs as a conditioned variant of the extremely common nasal sound *n* before velars ([k], [g], [x], [ɤ], etc.), juncture, etc. A related, minor complicating factor in determining the phonemic status of *ŋ*, in particular in those speech varieties where it may occur in initial position, is the presence in languages of prenasalized stops, where there is a homorganic nasal release before stop sounds, i.e. *ᵑg- ᵑk-*. Note that some languages, e.g. Swahili, a language of eastern Africa, actually contrast */ŋ/* and /ng/ (pronounced [ŋg]) in word-initial position (1).

(1) Swahili (Johnson and Madan 1991 [1939]: 335)

ŋoa	vs.	*ngoa*
'root up'		'passion, lust'

With regard to the phonotactics of phonemic *ŋ*, one finds an even more striking areal distribution across the world's languages. For example, while phonemic *ŋ* is found in all of the ten language families and isolate groups of Siberia, it is found word-initially only in those languages spoken in northern and eastern Siberia, e.g. Nganasan (Samoyedic, Uralic; north-central Siberia), Kerek (Chukotko-Kamchatkan; northeastern Siberia), Itelmen (Chukotko-Kamchatkan; eastern Siberia), Nivkh (isolate; southeastern Siberia and Sakhalin Island), Dolgan (Turkic; north-central Siberia), Kolyma Yukaghir (isolate; northeastern Siberia), and all Tungusic languages (a family spoken throughout central and eastern Siberia), but is lacking word-initially in Buriat (Mongolic; south-central Siberia), all Siberian Turkic languages except Dolgan (central Siberia), southern Samoyedic languages (Uralic; central Siberia), Khanty, Mansi (Ob-Ugric, Uralic; western Siberia), and Ket (Yeniseian; north-central Siberia). For more on *ŋ* in the languages of Siberia, see Anderson 2003a, 2003b.

(2) Nganasan (Tereščenko 1966a: 374)
ŋukagə 'many, a lot'
Kerek (Skorik 1986: 79, 85)
ŋuyŋən 'tail'
Itelmen (Volodin 1997: 63, 65)
ŋosx 'tail'
Nivkh (Gruzdeva 1998: 24)
ŋamk 'seven'
Dolgan (Ubrjatova 1985: 41)
ŋassa 'pipe'
Kolyma Yukaghir (Nikolaeva and Xelimskij 1997: 160)
ŋol- 'be'

Tungusic languages

Even (Novikova 1997: 291)
ŋɪn̄ 'dog'
Negidal (Cincius 1975: 663)
ŋoo- 'smell, sniff'
Nanai (Cincius 1975: 657)
ŋaala 'hand'
Ulcha (Cincius 1975: 668)
ŋele 'wolf' (< 'fear')

Some languages permit *ŋ* only in syllable-final or coda position. These include Burushaski (isolate; Pakistan or India), Hmong Njua (Hmong-Mien; southern China), Lower Grand Valley Dani (Trans-New Guinea; Papua, Indonesia), Mandarin, and West Greenlandic (Eskimo-Aleut).

(3) Burushaski (own field notes)
ariŋ 'my hand'
Hmong Njua (Lyman 1979: 10)
tóŋ 'edible gourd'
Lower Grand Valley Dani (Bromley 1961: 37)
saŋ 'bark container'
Mandarin (Li and Thompson 1981: 89)
fáŋzi 'house'
West Greenlandic (Vaxtin 1997: 98)
akiŋa 'answer me'

Other languages do not permit this sound in coda position at all. These include Fijian (Austronesian; Fiji), Gooniyandi (Bunaban; Western Australia), and Supyire (Gur, Niger-Congo; Mali), none of which permits any consonants in coda position. More interestingly, there are languages which lack *ŋ* in coda position, but do permit final consonants, e.g. Nenets (Northern Samoyed; Siberia), Margany (Pama-Nyungan; Queensland, Australia) and Canela-Krahô (Macro-Ge; Brazil).

(4) Fijian (Dixon 1988: 22)
ŋonevuli 'student'
Gooniyandi (McGregor 1990: 67)
ŋañi 'I'
Supyire (Carlson 1994: 26)
ŋɔ́ɔ́ 'sleep'
ŋágà 'scratch'
Nenets (Décsy 1966: 51)
ŋobkad 'often'
Margany (Breen 1981: 284)
ŋaɲbaɖ 'sweat'
Canela-Krahô (Shell 1952: 116)
ŋapar 'your nephew'

Less frequently, certain languages allow *ŋ only* in syllable-initial or onset position, for example Eastern Kayah Li (Karen, Tibeto-Burman; Myanmar).

(5) Eastern Kayah Li (Solnit 1997: 351)
ŋò 'language', 'weep'

2 Geographical distribution

In the basic 100-language sample of the *World Atlas of Language Structures*, 52 per cent lack this sound in their phonemic inventory

altogether. These include a diverse range of languages from all over the globe, for example Abkhaz (North-West Caucasian; Georgia), Asmat (Trans-New Guinea; Papua, Indonesia), Barasano (Tucanoan; Colombia), Basque, Chalcatongo Mixtec (Oto-Manguean; Mexico), Egyptian Arabic, Hausa (Afro-Asiatic; Nigeria), Khoekhoe (Central Khoisan; South Africa), Lakhota (Siouan; north-central United States), Modern Greek, Russian, and Turkish.

While many Papuan and New World (North and South American) languages lack the sound altogether, so that phonemic *ŋ* is rare or highly marked in these macro-areas, phonemic *ŋ* is found in such Papuan languages as Yimas and Enga, or in New World languages like the Yup'ik languages (Eskimo-Aleut; Alaska and northeastern Siberia), Cahuilla (Uto-Aztecan; southwestern California), Mapudungun (Araucanian; Chile and Argentina), or Lealao Chinantec (Oto-Manguean; Mexico).

(6) Yimas (Foley 1991: 39)
ŋarwa 'penis'
Enga (Lang 1973: 77)
ŋtli 'shot', 'nail'
Central Yup'ik (Jacobson 1995: 498)
ŋellar- 'to laugh'
Cahuilla (Bright 1965b: 243)
hé-naŋ 'his tongue'
Mapudungun (Salas 1992: 78)
ŋulu 'west'
Lealao Chinantec (Rupp 1989: 1)
ŋi^{H} 'his/her face'

In the New World languages, phonemic *ŋ* is most common in the south-west of North America (especially California), southern Mesoamerica and central Brazil. Phonemic *ŋ* is also quite common in Australian languages (where it is nearly universal), and in a band across central Africa. It is lacking in most Dravidian and Indo-European languages of South Asia, but is found in Burushaski, some Munda languages, and languages of the Tibeto-Burman family. Phonemic *ŋ* is quite common in northern Eurasian languages and nearly ubiquitous among the languages of South-East Asia. The distribution of *ŋ* in the Pacific, Australia, and South-East Asia has resulted primarily from the historical spread of the descendants of protolanguages having contrastive *ŋ*, including the protolanguages of the Austronesian, Tai-Kadai, Hmong-Mien, and Austro-Asiatic families of South-East Asia and the Pacific region, as well as Australian languages, regardless of whether or not these latter form a genealogical unit.

Thirty-two of the basic 100 languages allow *ŋ-* in word-initial position. These include such diverse languages as Bagirmi (Nilo-Saharan; Chad and Nigeria), Canela-Krahô (Macro-Ge; Maranhão, Brazil), Chukchi (Chukotko-Kamchatkan; northeastern Siberia), Indonesian, Mangarrayi (Australian; Northern Territory), Meithei (Tibeto-Burman; northeastern India), Oksapmin (isolate; Papua New Guinea), Rama (Chibchan; Nicaragua), and Vietnamese.

(7) Bagirmi (Stevenson 1969: 13)
ŋaŋ(a) 'tooth'
Canela-Krahô (Popjes and Popjes 1986: 161)
ŋũpar 'listen'
Chukchi (Volodin and Skorik 1997: 25)
ŋeekək 'daughter'
Indonesian (Echols and Shadily 1961: 230)
ŋelu 'have a headache'
Mangarrayi (Merlan 1982: 178)
ŋelŋel 'competent fisherman'
Meithei (Bhat and Ningomba 1997: 5)
ŋa 'fish'
Oksapmin (Lawrence 1993: 72)
ŋe 'rotted root of tree/reed'
Rama (Grinevald n.d.: 40)
ŋuu 'house'
ŋaaŋ 'bed'
Vietnamese (Nguyên 1966: 300)
ŋan 'swan'

Initial *ŋ-* is also found in Clallam (Salish; northwestern United States), in various indigenous languages of California, including Wikchamni (Penutian), Luiseño (Uto-Aztecan), and Washo (isolate), as well as in the Andamanese languages (e.g. Aka-Biada and Onge) spoken in the Indian Ocean (8).

(8) Clallam (Thompson et al. 1974: 187)
ŋəscən? 'louse'
Wikchamni Yokuts (Gamble 1978: 11)
ŋit^{h}eč^{h} 'snot'
Luiseño (Bright 1965a: 344)
ŋaaq 'he's weeping'
Washo (Kroeber 1907: 257)
ŋauŋaŋ 'baby'
Aka-Biada (Portman 1992 [1887]: 198)
ŋatia-da 'mangrove fruit sp.'
Onge (Dasgupta and Sharma 1982: 6)
ŋi 'you'

One language in this set, Grebo (Kru, Niger-Congo; Liberia), allows *ŋ* in initial position alone, but not in medial or final position (9).

(9) Grebo (Innes 1966: 35)
ŋe 'antelope'

Virtually all South-East Asian, Australian, and non-Papuan languages of the Pacific region have initial *ŋ-*. As mentioned above, the languages of northern and eastern Siberia also have initial *ŋ-*. Certain languages of New Guinea (e.g. Kâte (Finisterre-Huon; Papua New Guinea), Loniu (Austronesian; Papua New Guinea), Nabak (Finisterre-Huon, Trans-New Guinea; Papua New Guinea), Eipo (Mek, Trans-New Guinea; Papua, Indonesia), etc.) possess initial *ŋ-*, as well as a number of New World languages (Lummi (Salish; northwestern United States), Sochiapan Chinantec (Oto-Manguean; Mexico), Gavião (Tupian; Brazil)); but overall, *ŋ-* is not that common in these parts of the world.

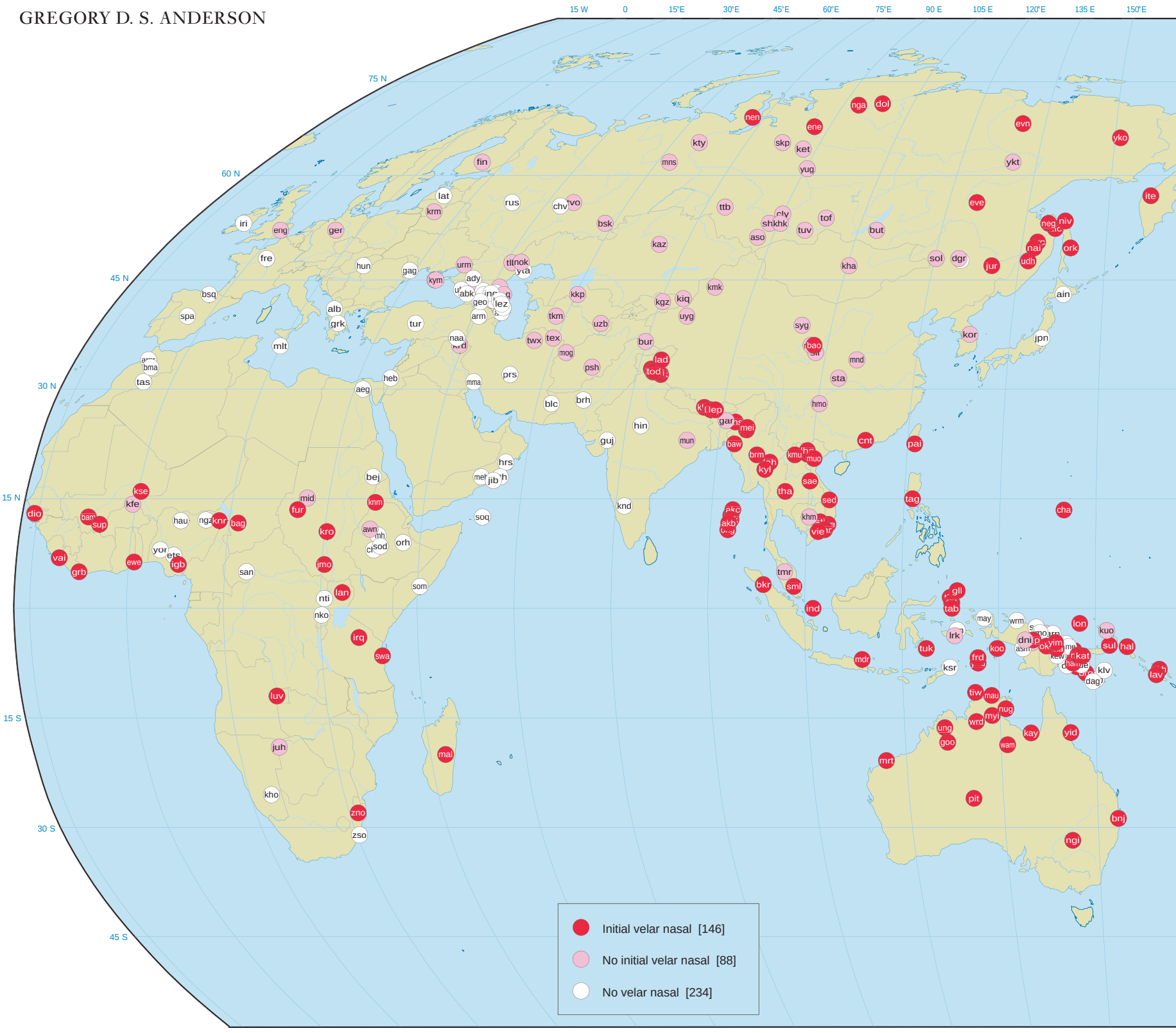
None of the handful of European and western Asian languages with phonemic *ŋ* permits this sound in word-initial position, nor do the languages of western and central Siberia, where phonemic *ŋ* is nearly universal. Thus, there is a large block of languages from Europe to central Siberia with phonemic *ŋ* which do not allow the sound word-initially. The majority of the Native American languages of California with phonemic *ŋ* do not permit the sound word-initially, while elsewhere (Africa, South America, Papua New Guinea) non-initial phonemic *ŋ* occurs only sporadically.

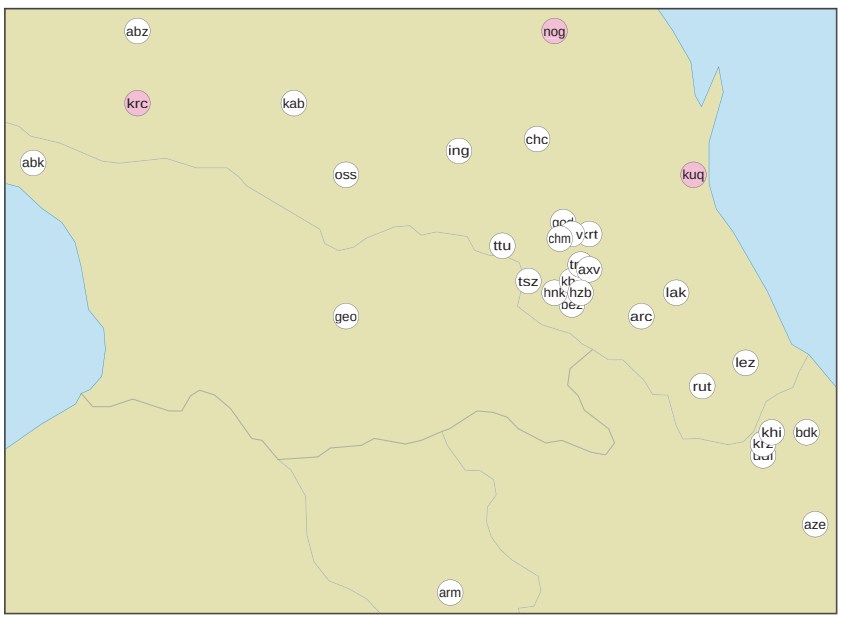
3 Theoretical issues

In the case of languages that allow phonemic *ŋ* in word-final or syllable coda position, but not in word-initial or syllable onset position, we find a violation of the well-known dictum, usually attributed first to Roman Jakobson (1941), that coda-position allows fewer contrasts or phonemes than onset position. Further, based on the phonotactic distribution of the velar nasal across the languages of the world, it appears that word-edge or word-peripheral and word-medial or word-internal syllable phonotactics are to be treated separately, and in particular, that word-edge coda and onset positions seem to be more restricted than corresponding coda and onset positions in non-edge positions. Lastly, while the macro-areal distribution and phonotactics of contrastive *ŋ* are relatively straightforward, there is actually considerable variation on the micro-areal level.

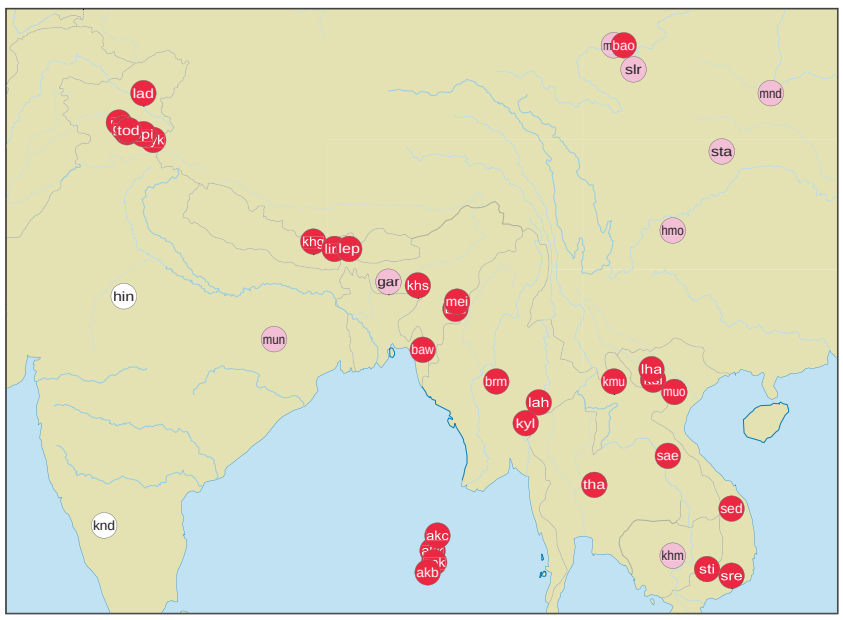
9 The Velar Nasal (*ŋ*)

GREGORY D. S. ANDERSON

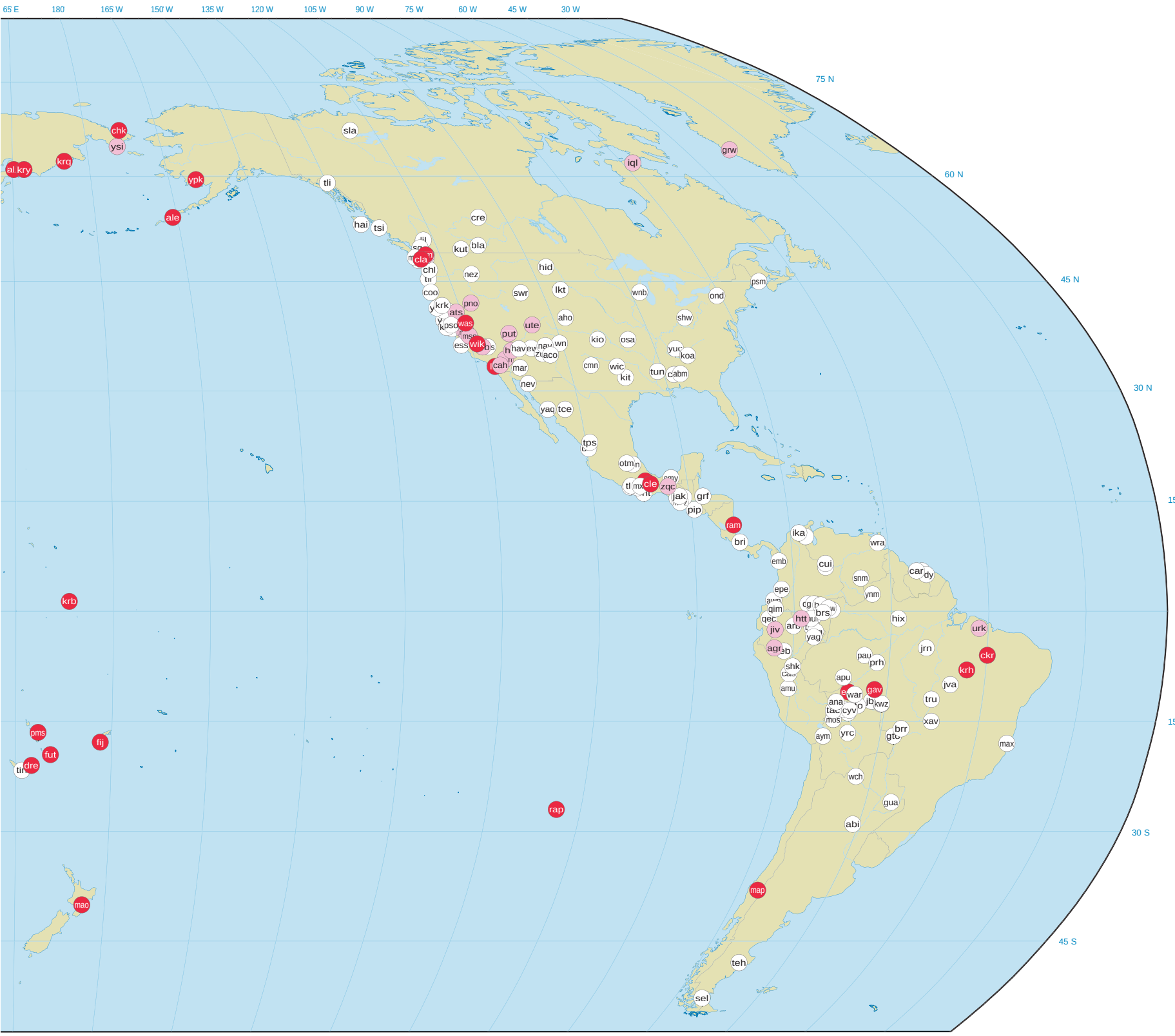




Caucasus



Kashmir to Vietnam



New Guinea and northern Australia

western United States

10 Vowel Nasalization

JOHN HAJEK

1 The values shown on the map

This map shows the distribution of contrastive vowel nasalization in the world's languages. The existence of contrastive nasal vowels is well known in some languages, e.g. French /pɛ̃/ *pain* 'bread' versus /pɛ/ *paix* 'peace'. About a quarter of the languages in our sample have a nasal–oral vowel contrast.

1.	Contrastive nasal vowels present	64
2.	Contrastive nasal vowels absent	179
	total	243

Nasalization phenomena are of course not restricted to oral versus nasal phonemic vowel contrast. Of the 179 languages with no nasal vowel contrast, at least 31 are reported to have some kind of non-phonemic nasalization, such as contextual nasalization of vowels adjacent to nasal consonants, e.g. Thai /mɛː/ → [mɛ̃ː] 'mother'.

2 Number of oral versus nasal vowels

The number of contrastive nasal vowels in a language is often less than that of oral vowels—approximately 60 per cent of our sample shows this pattern. Lakhota (Siouan; North and South Dakota) has five contrastive oral vowels, but only three nasal vowels:

(1) Lakhota (Boas and Deloria 1941)

i	u	ĩ	ũ
	o		
ɛ			
	a	ã	

There is evidence to suggest that a reduction in the number of nasal vowels has a phonetic basis: the acoustic effect of nasalization is known to affect the perceptual distinctiveness of vowels (e.g. Beddor 1993). However, it is difficult to predict the shape of reduced vowel inventories. Lakhota (1) has only maximally peripheral nasal vowels. In other languages, such as Yoruba (Niger-Congo; Nigeria), the distinction between close and open mid vowels is reduced in favour of nasal open mid vowels; but in Seneca (Iroquoian; New York State) the nasal vowel system is limited to mid vowels /ɛ̃ ɔ̃/ only:

(2) Yoruba (Ogunbowale 1970)

i	u	ĩ	ũ
e	o		
ɛ	ɔ	ɛ̃	ɔ̃
	a	ã	

Seneca (Chafe 1996)

i	u		
e	o		
		ɛ̃	ɔ̃
æ	ɑ		

Maba (Maban, Nilo-Saharan; Chad) is the most extreme example of reduced contrastivity: it has 12 oral vowels but only one nasal vowel /ũ/. In some languages nasal vowels are not only few in number, but are also extremely rare: there are only two recorded lexical items with nasal vowels in Karok (Hokan; California).

The very strong claim, made by Ferguson (1966), that "the number of nasal vowels is never greater than the number of non-nasal vowel phonemes" needs some slight revision. Koyra Chiini (Songhay; Mali), which has a four-way contrast involving nasality and length, has more short nasal vowels than short oral vowels:

(3) Koyra Chiini (Heath 1999b)

i	u	iː	uː	ĩ	ũ	ĩː	
e	o	eː	oː	ẽ	õ		õː
	a		aː	æ̃	ã		ãː

The reason for this unexpected imbalance is the borrowing of the front low vowel æ̃ in French loanwords, e.g. /maŋgazæ̃/ < French *magasin* 'warehouse'.

3 Vowel length and weight

Sixteen languages in the sample have a four-way contrast involving both long and short oral and nasal vowels. In many cases, the four-way system is fully symmetrical, as in Apurinã (Arawakan; Brazil):

(5) Apurinã (Facundes 2000)

i	ɨ	iː	ɨː	ĩ	ɨ̃	ĩː	ɨ̃ː
e	o	eː	oː	ẽ	õ	ẽː	õː
a		aː		ã		ãː	

The interaction of length and nasality means that nasal vowel systems can in fact be quite large. Matters are further complicated where Advanced Tongue Root is a third factor. Koromfe (Gur; Burkina Faso) has, as a result, 21 oral vowels and 19 nasal vowels, including ATR variants in both sets.

Hajek (1997) reports that vowel nasalization occurs preferentially on long vowels and bimoraic sequences. Such a claim finds additional support in Lango (Nilotic; Uganda), where nasal vowels only appear in a bimoraic sequence, e.g. /cɔŋ/ 'knee' versus /cõ.ẽ/ 'knees', /ʔorã.ã/ 'sterile'.

4. Nasal spreading

Vowel nasality has differing prosodic properties across languages. In many languages vowel nasalization is strictly local, i.e. it does not spread to an adjacent segment, even if vocalic, as in French /oseã/ which surfaces as [oseã] 'ocean'. In other languages, vowel nasality, inherent or from an adjacent nasal consonant, may spread onto and across adjacent segments. In Apurinã, all vowels adjacent to a nasal vowel are obligatorily nasalized, e.g. [ˈokɨ̃] 'eye of' + [ã] 'liquid of' = [oˈkɨ̃ã] 'eyeball'. Directionality of nasal spread across segments also varies. It may be mostly monodirectional as in Warao (Venezuela), where it is left to right, or bidirectional as in Barasano (Tucanoan; Colombia):

(6) a. Warao (Osborn 1966)

/inawaha/ → [inãwãhã]	'summer' (left to right)

b. Barasano (Jones and Jones 1991)

/to-hʉ/	'towards there'
/tĩbĩ-jã/ + /-hʉ/ → [tĩbĩ-jã-hʉ̃]	'to Otter River' (left to right)
/gahe/ 'other' + /-ĩ/ '3.M' = [gãhĩ]	'another male' (right to left)

There is a hierarchy of segmental permeability to nasal spreading (e.g. Cohn 1993). Nasality always spreads, when it occurs, through vowels, as in Apurinã. It typically also spreads through glides and through glottals, as in Seneca. It may also spread through less vocalic segments such as liquids and, more rarely, fricatives, as in Epena Pedee (Choco; Colombia):

(7) Epena Pedee (Harms 1994)

/wã/ 'go' + /-ɾú/ 'present' = [w̃ãn̆ú̃] 'go.PRES'
/sĩ́ɔso/ [s̃ĩ́ɔ̃s̃õ] 'sugar cane'

Guaraní (Tupian; Paraguay) is an extreme case of bidirectional spreading across all consonants, including voiceless stops, and blocked only by stressed oral syllables:

(8) Guaraní (Walker 1999)

/ⁿdo-ɾoi-ⁿdu-ˈpãi/ → [nõr̃õĩnũˈpãĩ]	'I don't beat you'
/aˌkãɾaˈuɥʷe/ → [ãˌkãr̃ãˈuɥʷe]	'hair (of the head)'

5 Vowel height and contextual nasalization

In most languages which show contextual nasalization (e.g. /NV/ → [Nṽ]), all vowels are affected equally regardless of vowel height. However, Hajek (1997) notes the existence of two very different patterns of interaction between vowel height and preferential nasalization. In some languages, there is evidence that high vowels are preferentially nasalized, e.g. Chamorro (Austronesian; Guam), whilst in others, e.g. Thai, low vowels appear to be preferred targets:

(9) a. Chamorro (Witucki 1974)
high: [dzæːmũ] 'you like' low: [mat.tʊ] 'arrived'

b. Thai (Noss 1964)
high: [mɯː] 'hand' low: [mãː] 'come'

6 Rhinoglottophilia

An association, sometimes termed "rhinoglottophilia", between vowel nasalization and glottal consonants has long been noted, e.g. by Matisoff (1975). In Thai there is predictable vowel nasalization of low vowels after word-initial /h/, e.g. /haː/ → [hãː] 'five'. In Pirahã (Mura; Brazil), there is optional nasalization after voiceless glottal stops and fricatives, e.g. [ʔapaí] ~ [ʔãpaí] 'head', [hói] ~ [hõi] 'one'.

7 Nasalization as word-boundary marker

In a small number of languages, nasalization serves as a word-boundary marker—evidence of such a phenomenon is found in Tashlhiyt (Berber; Morocco), where the oral vowels /i u a/ may be nasalized in word-final position:

(10) Tashlhiyt (Willms 1972)
nonfinal: [ˈmːᵘanːɛɣ] 'our mother'
final: [n̩ˈtːsã ʔɑdisːˈçɛnː ˌwã] 'it is he who made it'

8 Geographical distribution

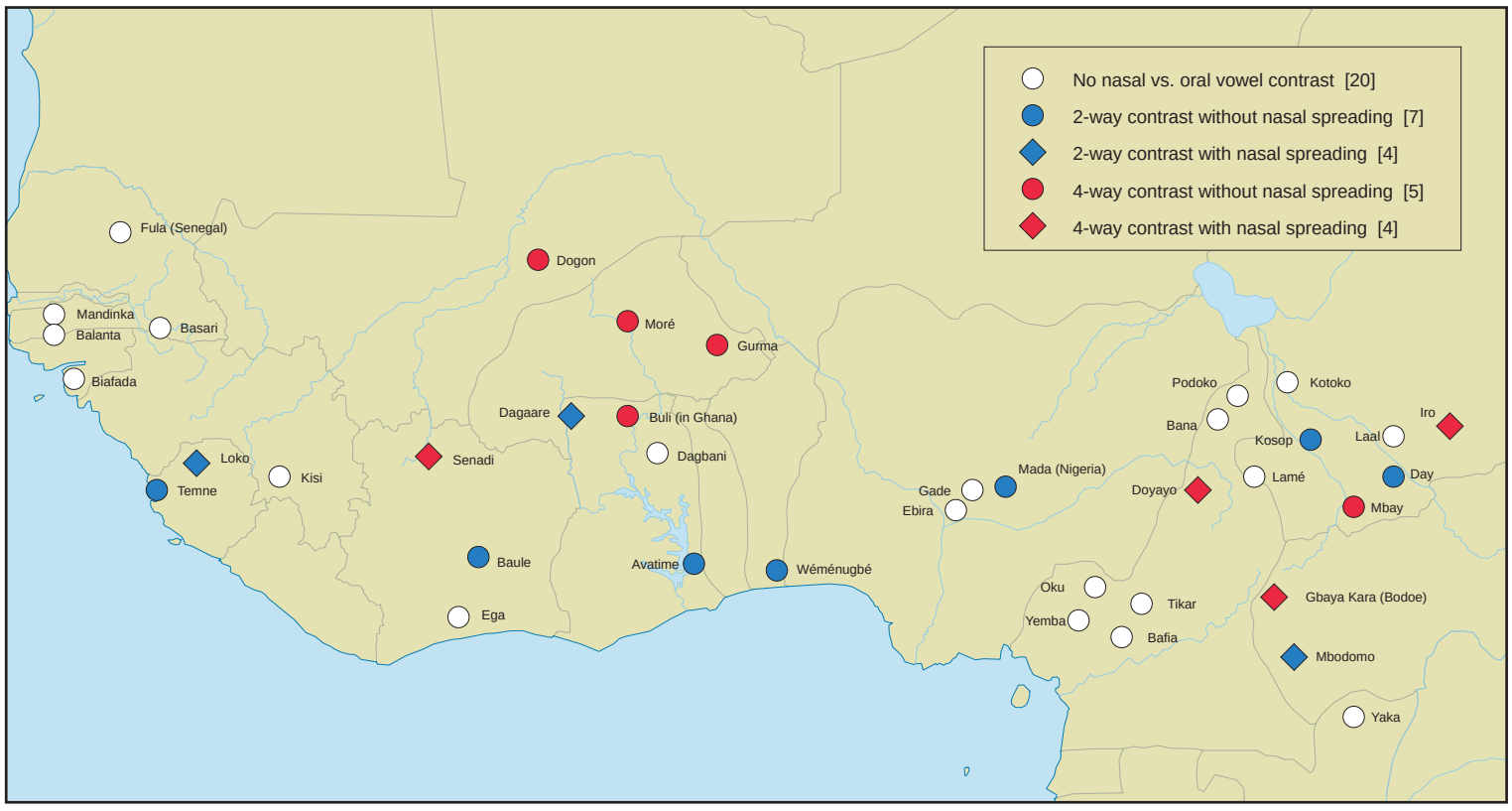
There appears to be strong areal influence in the distribution of contrastive nasal vowels around the world. As the map shows, they are concentrated in the following areas: the equatorial regions of South America and Africa, the northern part of the South Asian subcontinent, and throughout large parts of North America. Phonemically contrastive nasal vowels are, with few exceptions, otherwise absent from most of Eurasia, the northern and southern extremes of the Americas and Africa, as well as Australia and the Pacific.

9 Nasal vowels in West Africa

The inset map shows the complexity of nasal vowel systems in West Africa, an area already identified as having a high density of such systems. It distinguishes among languages that have no nasal versus oral contrast, those with only a two-way oral versus nasal contrast, and those that have larger four-way systems incorporating vowel length. It also indicates which languages have some kind of nasal spreading across segments.

1. No nasal vs. oral vowel contrast	20
2. Two-way nasal vs. oral vowel contrast (/ṽ/ vs. /V/) without nasal spreading	7
3. Two-way nasal vs. oral vowel contrast (/ṽ/ vs. /V/) with nasal spreading	4
4. Four-way nasal vs. oral vowel contrast (/ṽ/ vs. /ṽː/ vs. /V/ vs. /Vː/) without nasal spreading	5
5. Four-way nasal vs. oral vowel contrast (/ṽ/ vs. /ṽː/ vs. /V/ vs. /Vː/) with nasal spreading	4
total	40

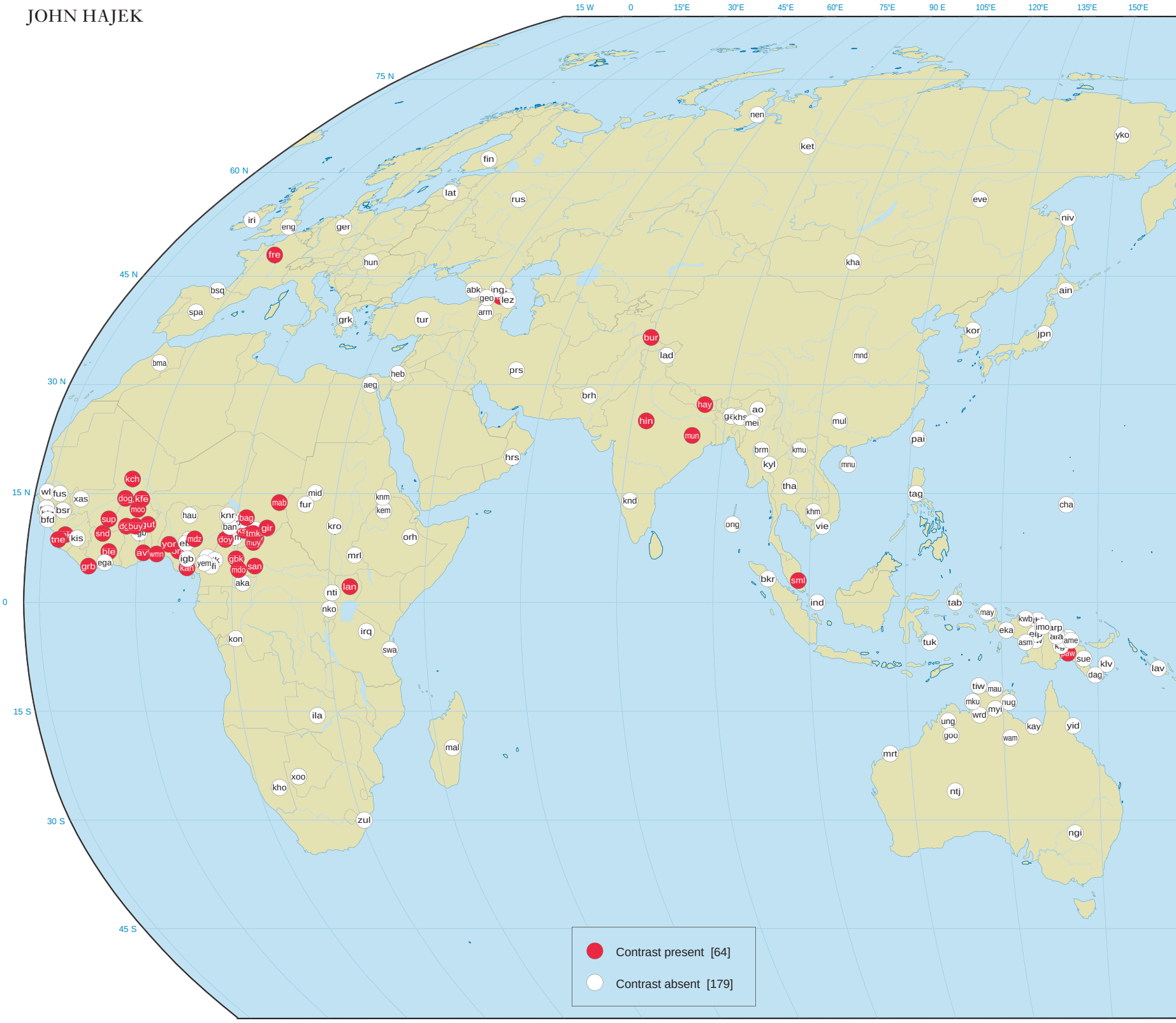
Values for Map 10A Nasal Vowels in West Africa

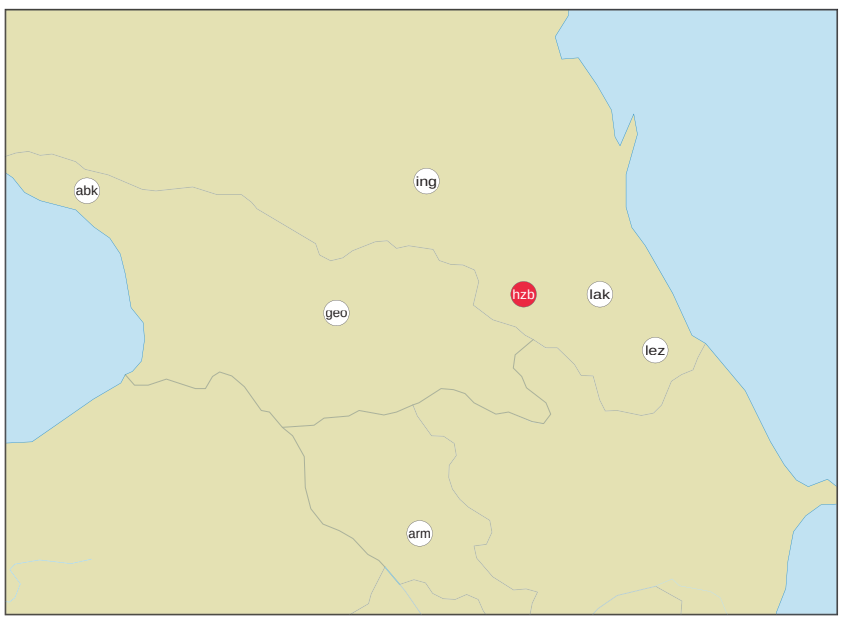


Map 10A Nasal Vowels in West Africa

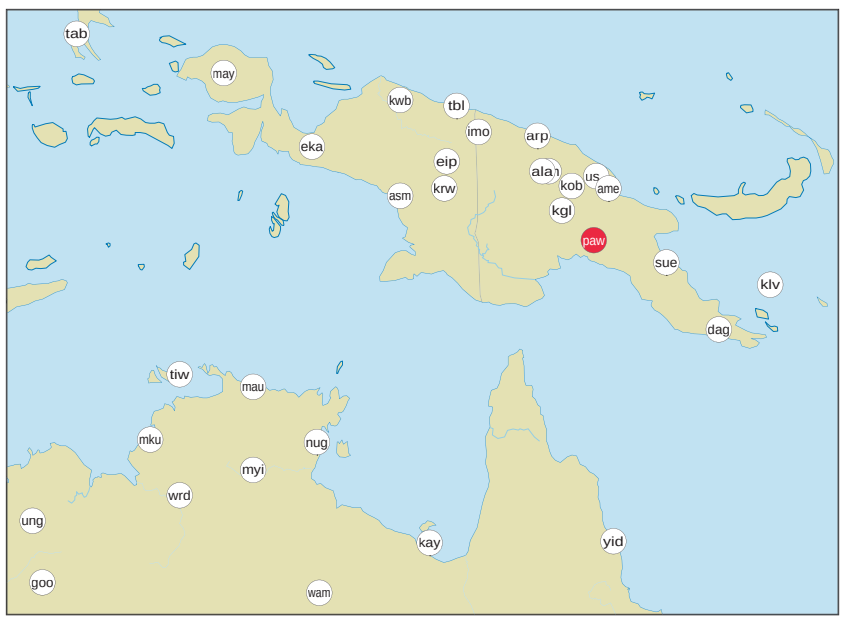
10 Vowel Nasalization

JOHN HAJEK

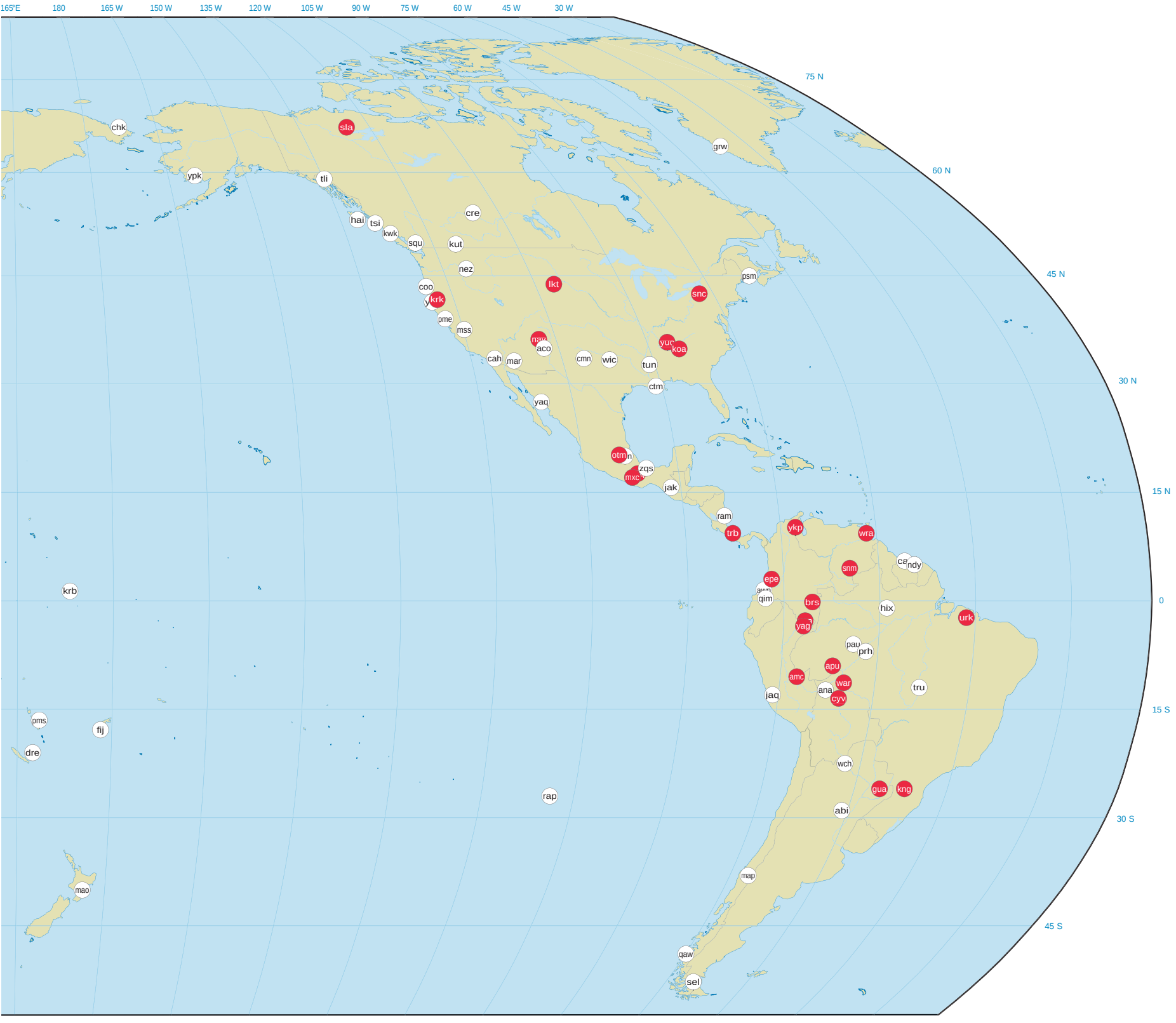


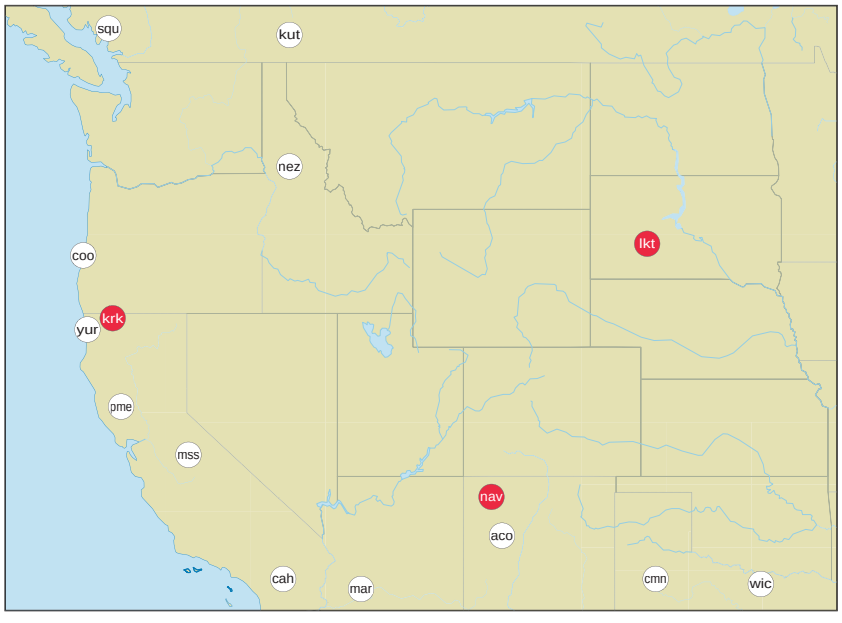


Caucasus

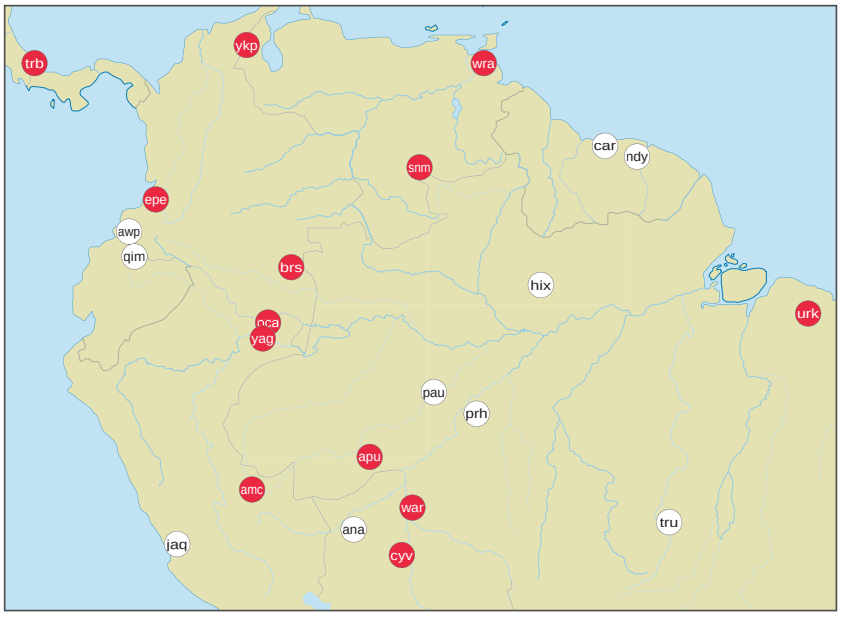


New Guinea and northern Australia





western North America



northwestern South America

11 Front Rounded Vowels

IAN MADDIESON

1 Introduction

The lip position of vowels is generally correlated with the height and backness of the vowel, as discussed in Chapter 2. Normally vowels which are back and not low are pronounced with the lips in a rounded position (e.g. [o] and [u]). Front and central vowels and back low vowels are normally pronounced with the lips unrounded (e.g. [i], [e], [a]). Lip rounding involves drawing the corners of the lips together and protruding the lips forward from their normal rest position. The ability to make these gestures with the lips is greatest when the vowel is high; as the jaw is opened further to make progressively lower vowels, the amount of adjustment of the lips that is possible becomes more limited, since the lips are being stretched vertically. This mechanical constraint may account for the fact that low vowels are normally unrounded, whether front or back. As for non-low vowels, the association of lip rounding with back vowels has to do with reinforcing the low-pitched acoustic overtones that are characteristic of back vowels. Rounding and protruding the lips lowers the frequencies of the acoustic resonances of the vocal tract to add to the lowering effect produced by the action of backing the body of the tongue. A rounded back vowel is therefore more clearly distinct from other vowels than an unrounded one (see Stevens 1972).

Although the occurrence of lip rounding with back non-low vowels is the normal pattern, front vowels with lip rounding and back non-low vowels without lip rounding do both occur in the inventories of vowels of a minority of languages. This chapter will focus on the occurrence of one class of these less usual vowels, the **front rounded vowels**. In the sample of languages surveyed for this chapter, there are only 37 languages (about 6.6 per cent) in which one or more such vowels are recognized as part of the vowel inventory as defined in Chapter 2.

1. None	524
2. High and mid	23
3. High only	8
4. Mid only	6
total	561

Front rounded vowels typically occur in vowel systems with a larger than average number of vowel qualities. The mean number of vowel qualities in the 37 languages concerned is 7.87, whereas the mean for all the languages without front rounded vowels is 5.85, a difference which is statistically highly significant. In most cases when any front rounded vowel occurs, the inventory of vowels has a front unrounded vowel of similar height, and also a back rounded vowel of similar height. For example, French, Albanian, Mandarin, Dagur (Mongolian; China), and Lezgian (Nakh-Daghestanian; eastern Caucasus) have a high front rounded vowel, the phonetic symbol for which is [y]. This symbol is chosen following the spelling tradition of languages such as Norwegian and Finnish, which also have this sound. The vowel systems of all these languages also include both a high front unrounded vowel ([i]) and a high back rounded vowel ([u]). Because of these facts, as well as what is known about the history of a good many of the individual languages concerned, the presence of front rounded vowels is generally taken to reflect an elaboration of a vowel system that had fewer members at an earlier time.

2 Geographical distribution

The majority of the languages with front rounded vowels, 29 of the 37, are found in the more northerly parts of the Eurasian landmass. Outside of this area the few languages reported with front rounded vowels are widely scattered, and the reports are not always sufficiently detailed to be relied on with confidence. Only four languages of the Americas surveyed are reported to have one or more front rounded vowels: Hopi (Uto-Aztecan; Arizona), Quiotepec Chinantec (Oto-Manguean; southern Mexico), Wari' (Chapacura-Wanhan; Brazil), and Aikaná (Arawakan; Brazil). Only one of the sampled languages in Africa, Ejagham (Bantoid; Nigeria and Cameroon), has been reported as having a front rounded vowel. One other African language in the survey, Wolof (Atlantic, Niger-Congo; Senegal and Gambia), has sometimes been said to have a mid front rounded vowel, but in fact this vowel seems to be central and unrounded. The report of a front rounded vowel may be because of the similarity to the ear of the two types of sounds. Given that most linguists working in Senegal, where Wolof is spoken, are familiar with French, it seems that they have described this vowel as having the characteristics of the most similar-sounding vowel of French. In the Pacific region there are just three languages in the survey with front rounded vowels: Malakmalak (Daly; Northern Territory, Australia), Iaai (Oceanic; Ouvéa, New Caledonia), and Nambakaengö (Reef Islands-Santa Cruz family; Solomon Islands). Among these scattered languages the occurrence of front rounded vowels is certain in Iaai and Wari'. For the remainder it is possible that an auditory confusion between front rounded and central unrounded vowels might account for the suggestion that these languages have front rounded vowels.

Accepting all the data discussed above as valid, the majority of the languages in the survey with front rounded vowels have them in both the high and mid vowel classes; this is the case for 23 of the 37. These languages are placed in one class. Of the remaining languages, eight have only high vowels in their front rounded set and six have only mid vowels in their front rounded set. Altogether there are thus 31 languages in the survey whose vowel inventories include high front rounded vowels and a total of 29 languages whose vowel inventories include mid front rounded vowels. These numbers are so similar that it is not possible to talk of any preference for front rounded vowels to be high rather than mid. The strong pattern is for both high and mid to occur together in the inventory of vowels.

3 Discussion

Although they are certainly not present in all of the languages in the area, it is clear that front rounded vowels are most "at home" in the languages of Europe and Asia, apart generally from South and South-East Asia. They occur in a good proportion of the Uralic and Altaic languages, including Finnish, Hungarian, Selkup, and Nganasan in the Uralic group and Turkish, Yakut, Dagur, and Manchu in the Altaic group. These languages often have a pattern of vowel harmony according to which all the vowels in a word are required to agree in some of their properties; for example, all vowels might be required to be front. Thus, in Hungarian one of the case suffixes on nouns, traditionally called the essive, has the shape /-ul/ containing a high back rounded vowel if the noun stem has back vowels, but is /-yl/ (orthographic *ül*) if the stem has front vowels. In Turkish, in addition to agreeing in frontness, vowels in certain situations must agree in rounding with a preceding rounded vowel. It seems likely that the historical processes that bring front or rounded vowel harmony into the sound system of a language might be instrumental in introducing the front rounded vowels into the inventory. (It should be noted that vowel harmony patterns in individual languages are often more complex than the simple outline given here; for some discussion see the contributions in Vago 1980.)

Front rounded vowels can also be developed independently of vowel harmony, however. French is among the languages whose high front rounded vowel [y] arose as a consequence of an earlier high

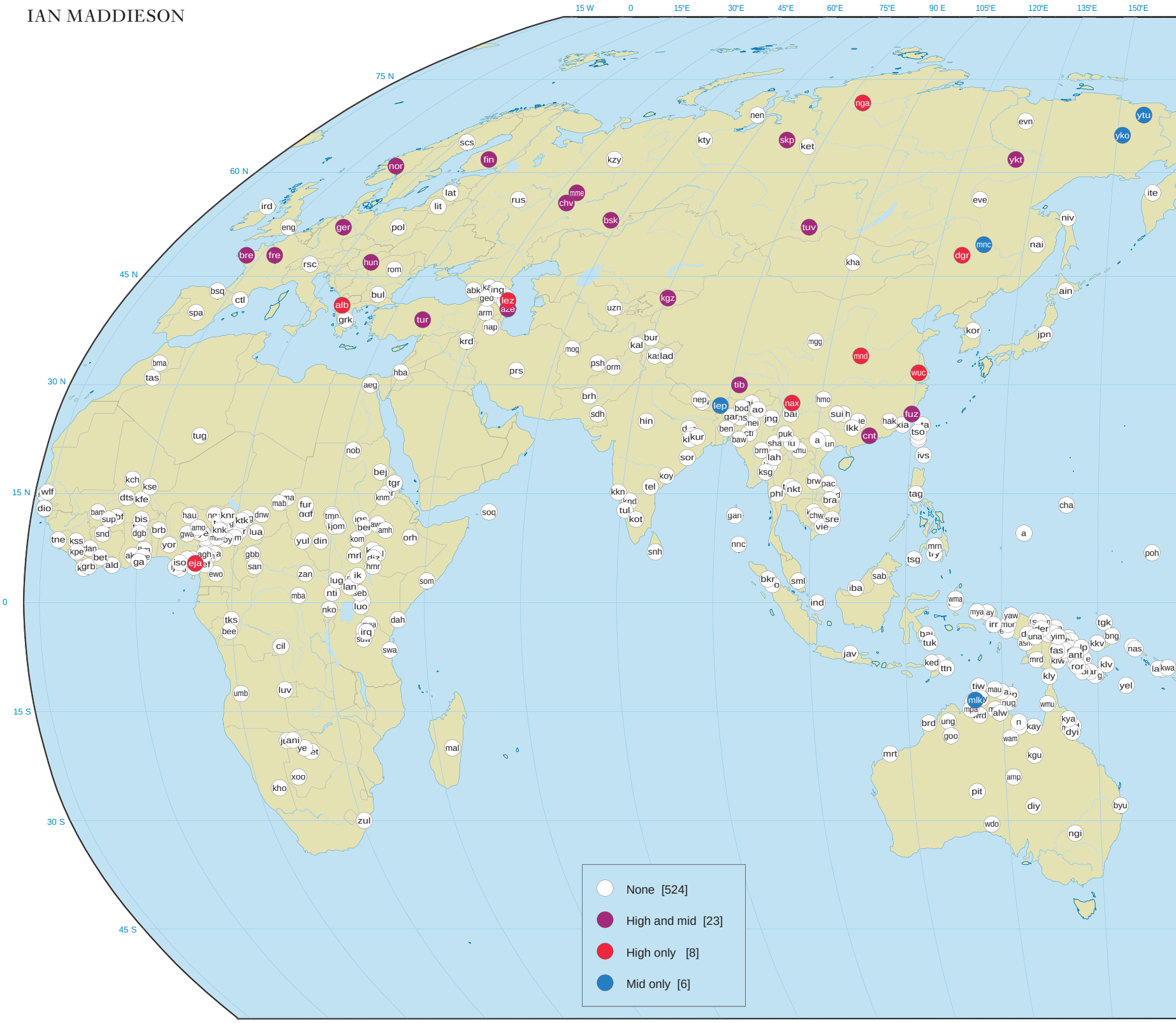
back rounded vowel [u] moving forward while keeping the rounding. In French a high back rounded vowel [u] has been "re-created" in the system by mid vowels rising in height. For example, the familiar second-person subject pronoun *tu* is now pronounced [ty]—it is still spelled the same as its Latin source *tu* but in Latin pronunciation the word had a back vowel. The plural or polite second-person pronoun *vous* is pronounced [vu] or [vuz] (depending on context), with a high back vowel, but this derives from the Latin *vos* with a mid vowel. A number of varieties of modern English, such as those spoken in Australia, are well on the road to completing a similar change of the formerly back vowel, in words such as *food*, into a front rounded vowel. In these cases the creation of front rounded vowels is not conditioned by adjacent sounds in the word.

In other languages front rounded vowels have evolved from the transfer of the front property of an adjacent segment to a once-back vowel. A well-known example is German, where at some time in the Middle Ages back vowels in many words, when followed by a high front vowel, were changed into front vowels while keeping their other properties. The vowels causing this change were frequently in derivational suffixes or inflectional endings, and were subsequently lost or changed into mid vowels. The result is many pairs of words such as the singular/plural noun forms written *Kuh* 'cow', *Kühe* 'cows', in which the first has the vowel [u] and the second its front counterpart [y] followed by a mid vowel. Front rounded vowels have also evolved in some languages from the influence of a front consonant on a back rounded vowel. In the history of the Chinese languages one of the sources of the [y] vowel is a sequence such as /ut/. The tongue has to be moved forward to make the contact for /t/; consequently a more fronted position of the vowel is created and this pronunciation of the vowel becomes the new norm over a period of time. If the final consonant is later lost a change from [ut] to [y] is completed.

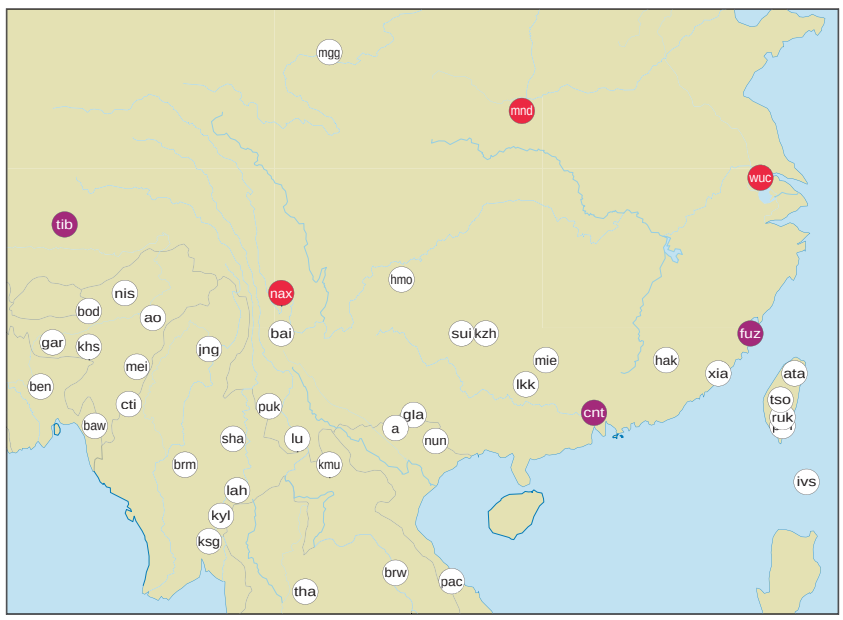
In view of the different historical scenarios which produced the front rounded vowels in various languages, it is quite striking that their occurrence is so relatively concentrated in a particular geographical area. It seems likely that the hearing of sounds of this sort in some languages of the area may have given further support to phonetically natural processes in other languages, with the end result being the addition of front rounded vowels to the inventory of more of the languages.

11 Front Rounded Vowels

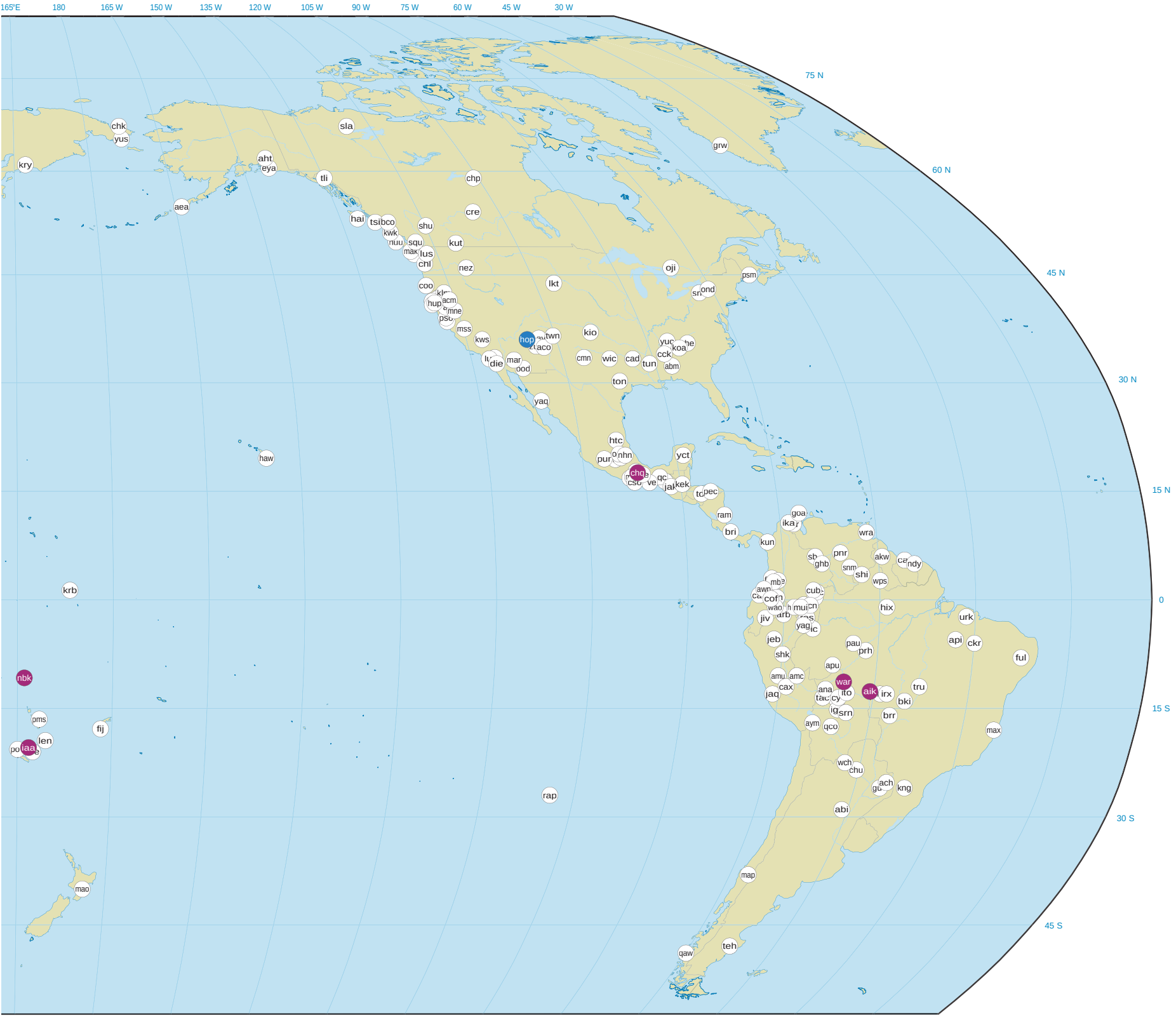
IAN MADDIESON

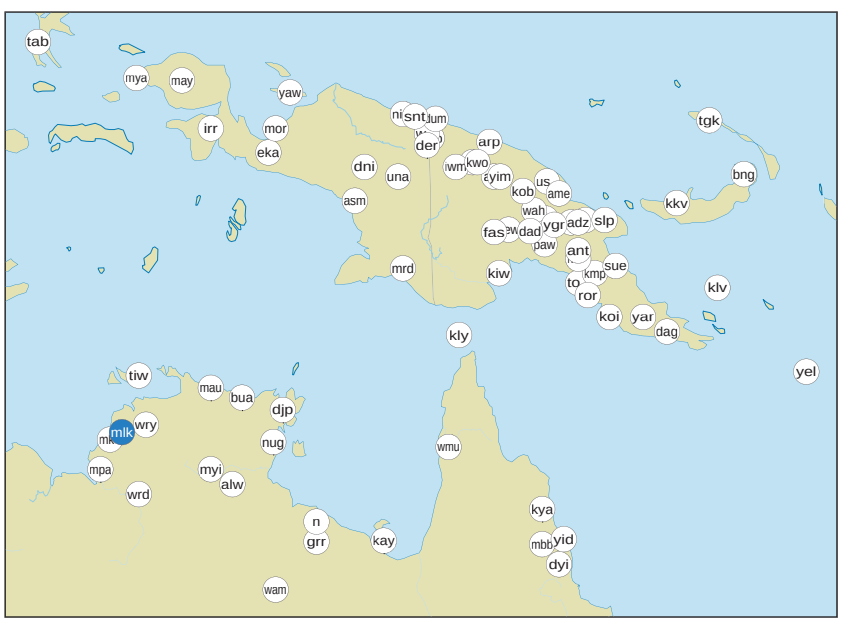


Caucasus and Black Sea area

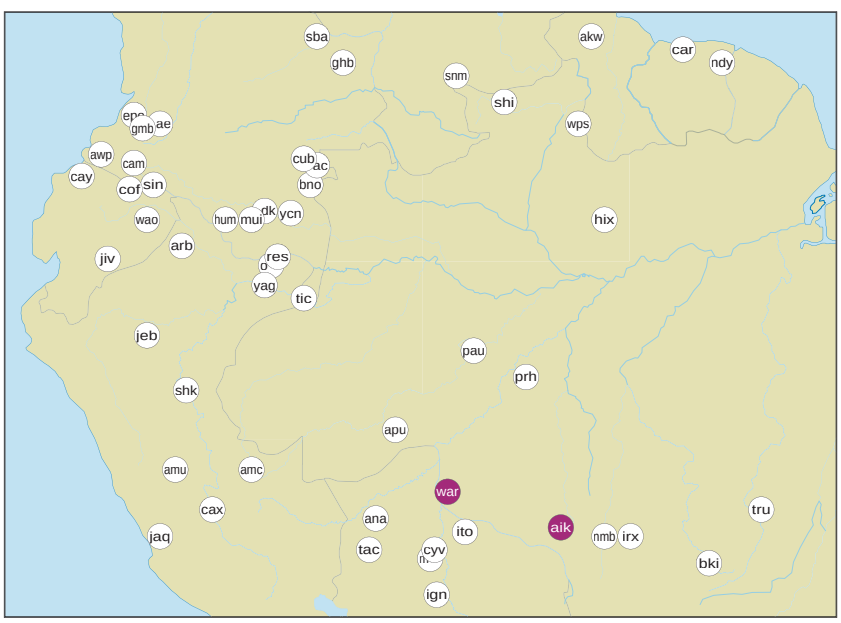


southern China and adjacent areas





New Guinea and northern Australia



northwestern South America

12 Syllable Structure

IAN MADDIESON

1 Introduction

The size of the sets of consonants and vowels which form the segment inventories of languages has been discussed in Chapters 1 and 2. In addition to the number of segments that languages use, it is also important to consider the ways that the segments are allowed to combine with each other in making longer structures, such as words and syllables. Some languages allow very free combination of segments, while in others the combinations are strongly restricted. In this chapter the complexity of sequencing of segments within syllables will be discussed as a means of examining one important aspect of how the combination of individual sounds is governed across the sampled set of languages.

The **syllable** is a well-recognized unit in linguistic analysis which explains quite well the number of rhythmic units that will be perceived in a word or longer utterance. This number is usually equal to the number of vowels in the utterance. Although it is usually easy to get agreement on the number of syllables present in a word, intuitions sometimes differ over where the boundaries between one syllable and another should be placed. Despite such hesitation, the syllable has proven to be a very useful concept in discussing the general rules for distribution of sounds in languages. Where listeners differ in syllabifying particular words, it is generally the case that both possible syllabifications can be shown to be permitted ones since unambiguous cases of each type can be found. For example, an English word such as *pastry* might be syllabified by different speakers as *past.ry* or as *pas.try* (where the dot represents a division between syllables). Since both *paste* and *tree* are perfectly acceptable monosyllabic words of English, either division will agree with a broader rule concerning possible syllables of English. The broadest rules of this kind for any given language describe what is called the canonical syllable pattern of the language. This is the pattern which essentially characterizes how many consonants may occur before the vowel in a syllable, and how many after the vowel.

2 Defining the values

Canonical syllable patterns are most often represented as a string of C and V symbols, where C stands for a Consonant and V for a Vowel sound (including any complex vowel elements such as diphthongs which might occur in the language). The one kind of syllable which seems to occur in every language is CV, that is, a syllable consisting of just one consonant preceding a vowel. In a relatively small number of languages this is the only type of syllable permitted. Such languages include Hawaiian and Mba (Adamawa-Ubangian, Niger-Congo; Democratic Republic of Congo). It is more frequent to find languages in which it is permitted not to have an initial consonant, as for example in Fijian, Igbo (Niger-Congo; Nigeria), and Yareba (Yareban; Papua New Guinea). For these languages the canonical syllable can be represented as (C)V, the parentheses indicating that an initial consonant is an optional element. If a language only allows syllables which fit this template, the language will be said to have **simple syllable structure**.

A slightly more elaborate syllable structure would add another consonant, either in the final position of the syllable or at its beginning, giving the structures CVC and CCV; these are both modest expansions of the simple CV syllable type. But it is worthwhile to make a distinction between two types of two-consonant strings. In a very large number of languages, although two consonants are allowed in the onset position of a syllable, there are strict limits on what kinds of combinations are permitted. The second of two consonants is commonly limited to being one of a small set belonging to either the class of "liquids" or the class of "glides". The liquids are the sounds commonly represented by the letters *r* and *l*, while glides are vowel-like consonants such as those at the beginning of the English words *wet* and *yet*. Liquids and glides have in common that they are produced with a configuration of the speech organs which permits a relatively unobstructed flow of air out of the mouth. Languages which permit a single consonant after the vowel and/or allow two consonants to occur before the vowel, but adhere to a limitation to only the common two-consonant patterns described above, are counted as having **moderately complex syllable structure**. An example is Darai (Indo-Aryan; Nepal). Here the most elaborate syllable permitted is CCVC, as in /bwak/ '(his) father', but the only possible second consonant in a sequence of two is /w/.

Languages which permit freer combinations of two consonants in the position before a vowel, or which allow three or more consonants in this onset position, and/or two or more consonants in the position after the vowel, are classified as having **complex syllable structure**. An obvious example of complex structure is English, whose canonical syllable pattern is often cited as (C)(C)(C)V(C)(C)(C)(C). The full expansion of the pattern only occurs in a few words such as *strengths*, when pronounced /stɹɛŋkθs/, but it is relatively easy to find syllables beginning with three consonants or ending with four, as in *split* and *texts* (/tɛksts/).

The classification of languages into three categories of syllabic complexity, simple, moderate, and complex, naturally overlooks many other questions of segment distribution (for example, whether the syllables at the beginnings and ends of words have the same or different restrictions from those which are internal to words), and has to gloss over some important differences with respect to how rarely or frequently the more complex syllable types occur in a given language. In reaching decisions regarding how to classify a given language, a certain common-sense flexibility has been employed. For example, if some kinds of consonant sequences have only recently been introduced into a language as a result of borrowing international words (such as *sport* or *golf*) the language will be classed on the basis of what occurs in more established vocabulary. Despite its summary nature the three-way classification provides a useful grouping with interesting geographical characteristics.

○	1. Simple syllable structure	61
○	2. Moderately complex syllable structure	274
●	3. Complex syllable structure	150
	total	485

3 Geographical distribution

By far the most common type are languages which permit moderately complex syllable structure, accounting for about 56.5 per cent of the sample. This type is of course widespread, but it is particularly frequent in Africa, the more easterly part of Asia, and much of Australia. Only 61 (about 12.5 per cent) of the languages in the sample limit themselves to nothing more complicated than the universal syllable type, CV. 150 languages (about 30.9 per cent of the total) permit complex syllables of one or more types. The languages with simple syllable structures tend to be distributed somewhere near the equator, in Africa, New Guinea, and South America. Note that this distribution has a considerable overlap with that found for languages with smaller consonant inventories, with the fit being least good in Africa. Languages with complex syllable structures are predominantly found in the northern two-thirds of the northern hemisphere, that is, in northern North America and northern Eurasia, where this type actually dominates over the others. Here there is a degree of overlap with the areas in which large consonant inventories tend to be more frequent, most especially in the more northerly parts of the North American continent. A smaller cluster of languages with complex syllable structure is found in northern Australia.

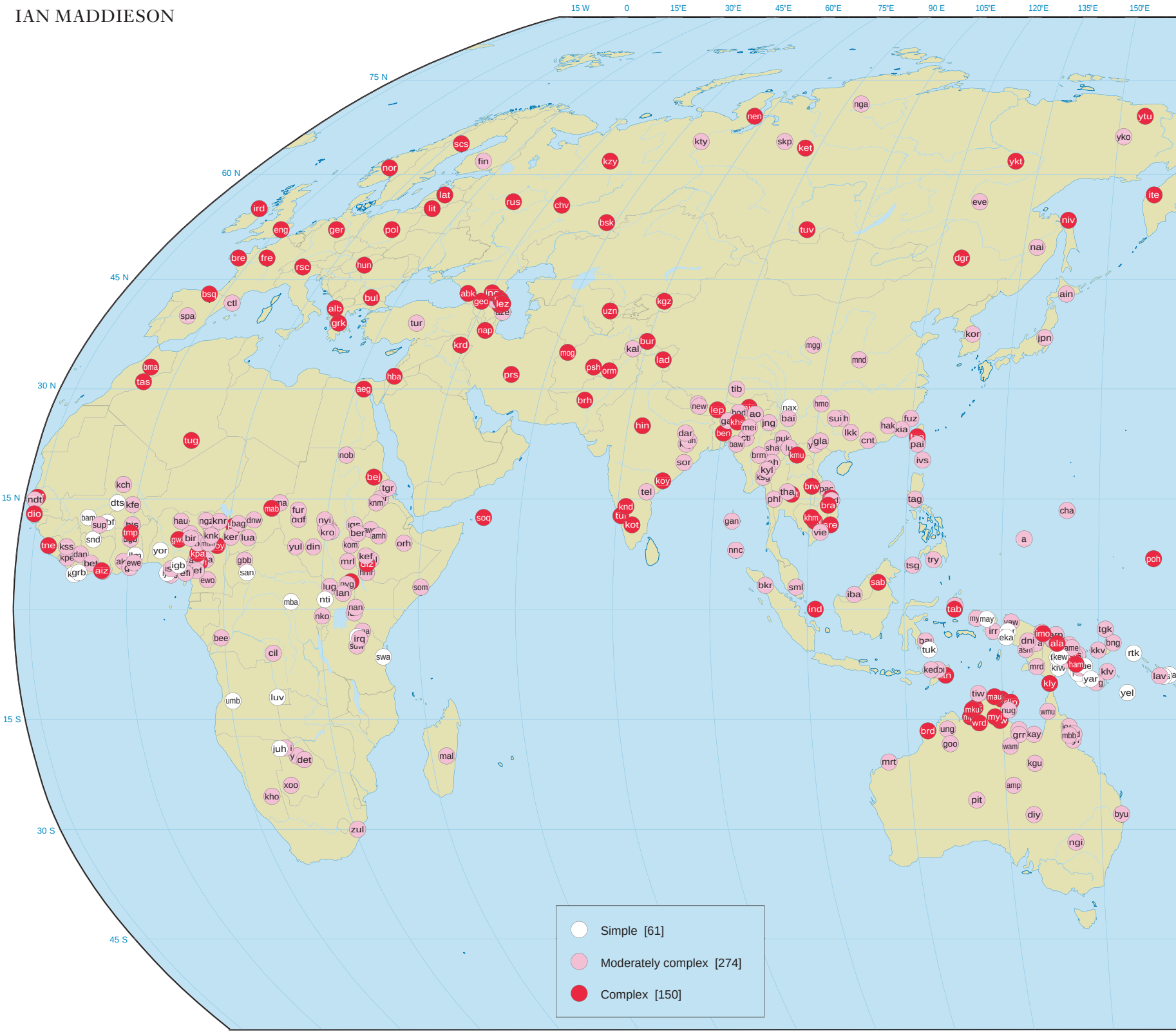
4 Correlations

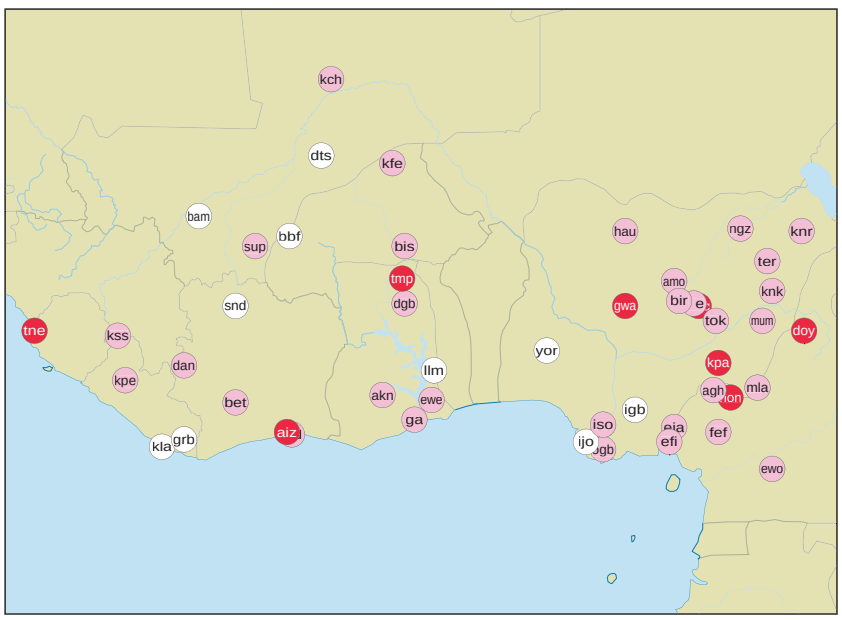
The areal overlaps noted above between small consonant inventories and simple syllable structure, and large consonant inventories and complex syllable structures, provide an interesting example of how complexity in different areas of phonological structure may work together as mutual reinforcement rather than being mutually offsetting. Across the set of 484 languages for which both consonant inventory size and syllable structure data are included, there is a significant though not strong correlation between these two measures of complexity (the correlation coefficient is .203 and the statistical significance level is very high, $p < .0001$). The languages with simple canonical syllable structure have an average of 19.1 consonants in their inventory, languages with moderately complex syllable structure have an average of 22.0 consonants, and those with complex syllable structures have an average of 25.8 consonants. As discussed in Chapter 13, syllable structure complexity does not show a similar pattern of positive correlation with tone system complexity.

It should be noted that the patterns of association between syllable structure and consonant inventory size established here are assuredly, at least in part, a reflection of the overlapping geographical distributions of properties noted in §3, and these overlaps might be due to fortuitous distribution of genealogically shared or areally spread features. Analysis of the patterns within and between different linguistic areas and language families would be required to decide whether the association should be attributed to accidents of survival and spread of particular languages, or can be proposed as reflecting a design feature of language viewed as a whole.

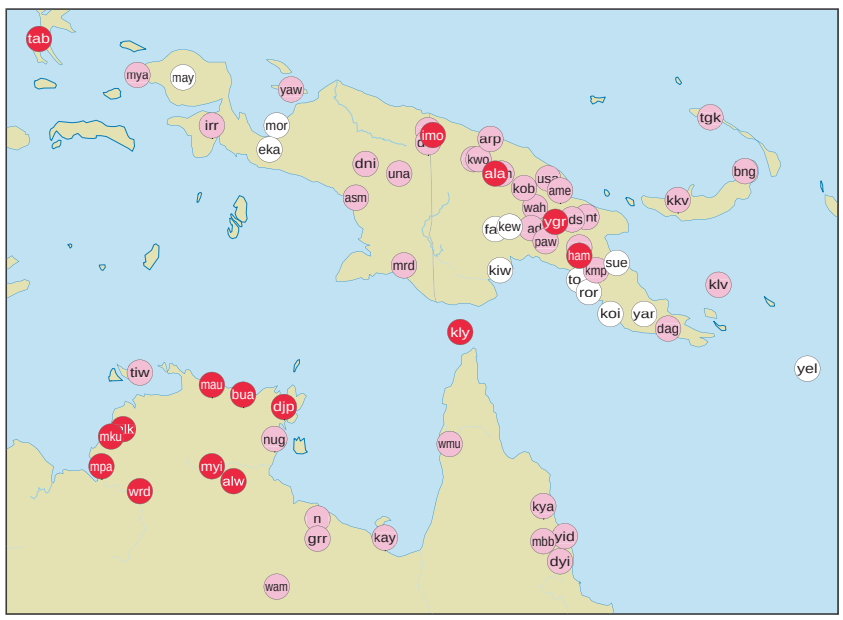
12 Syllable Structure

IAN MADDIESON

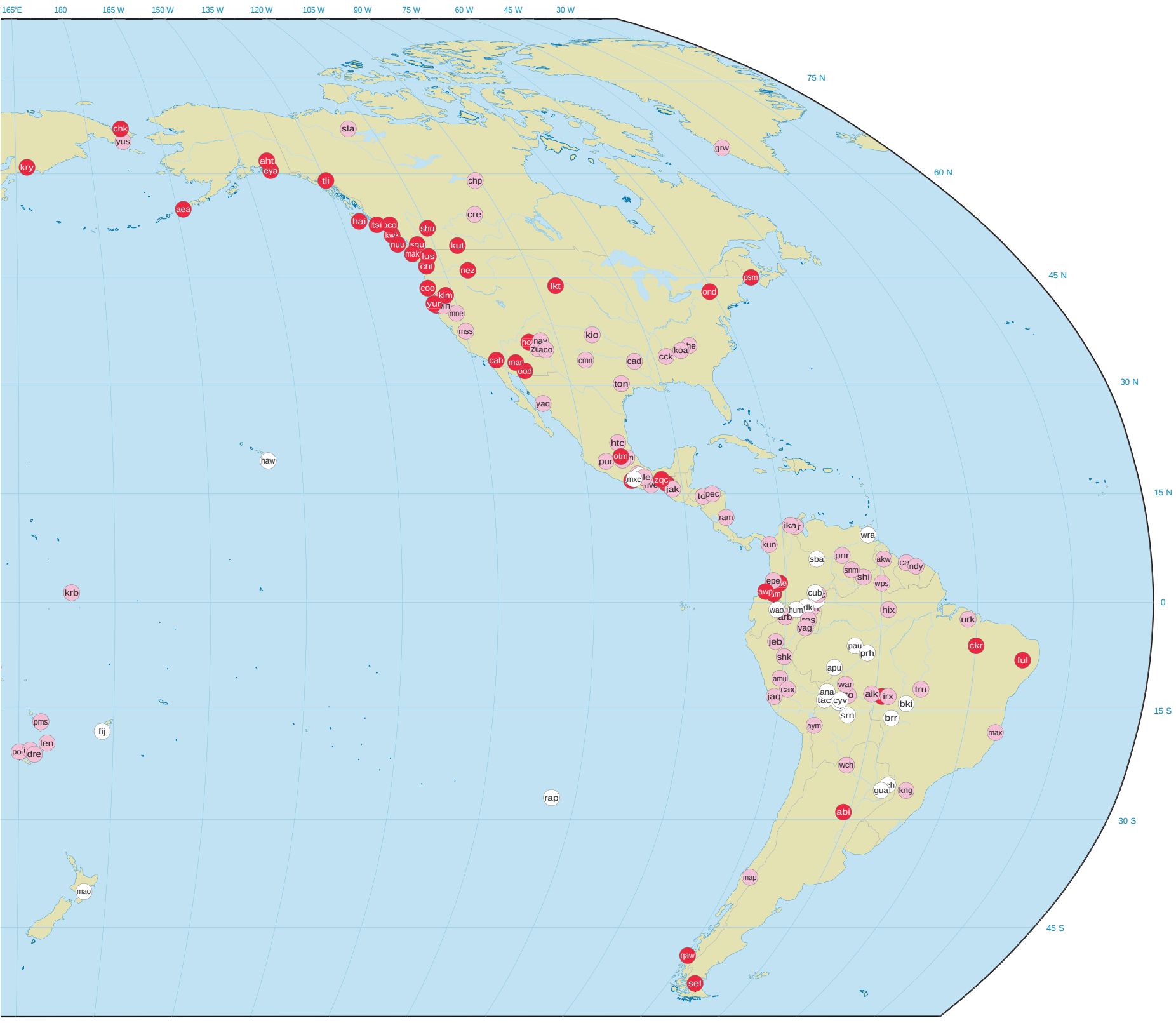




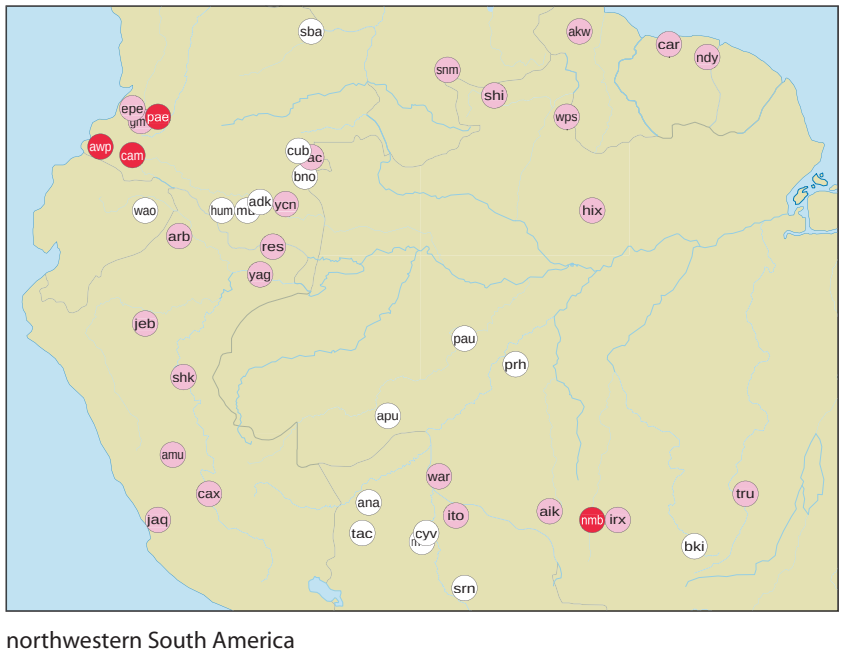
West Africa



New Guinea and northern Australia



Mesoamerica



northwestern South America

13 Tone

IAN MADDIESON

1 Introduction

All languages make use of variations in the musical pitch of the voice as part of their sound systems, but they differ in the ways in which modifications of pitch are used and how many different types of functions are served by pitch variations. Linguists distinguish between two of the major uses of pitch as *tone* and *intonation*. **Intonation** is the term that is used to describe sentence types, such as question versus statement, or to indicate whether a speaker has finished or intends to continue speaking, or to show which parts of an utterance present new or highlighted information versus old or less significant information.

Tone is the term used to describe the use of pitch patterns to distinguish individual words or the grammatical forms of words, such as the singular and plural forms of nouns or different tenses of verbs. In the simplest cases, each syllable of a language with tones will have its own characteristic tonal pattern, which may be a relatively flat pitch at a particular level, or may involve the pitch rising or falling over the duration of the syllable. When the pitch has a moving pattern of this sort, the tone is described as a **contour tone**.

Many of the languages of East and South-East Asia, including all varieties of Chinese as well as Vietnamese and Thai, have tone systems that include contour tones. In Thai, for example, full syllables have one of five tones, of which three are basically flat at high, mid, and low pitch levels, and the remaining two are rising and falling in shape. The syllable shape /kʰaa/ thus corresponds to five different words depending on the tone—/kʰáá/ with high tone means 'to trade', /kʰāā/ with mid tone means 'to get stuck', /kʰàà/ with low tone means 'galangal', /kʰàá/ with rising tone means 'leg', and /kʰáà/ with falling tone is a declarative politeness particle used by female speakers. Contour tones occur in other parts of the world, particularly in Mexico and Central America, but languages in which it is only necessary to speak of level tones are more common. Yoruba (Defoid, Niger-Congo; Nigeria) has lexical distinctions between three level tones, as in the three verbs /bí/ 'give birth', /bī/ 'ask', /bì/ 'vomit', while Navajo (Athapaskan; Arizona, etc.) has a contrast between high and low levels, as in the words /t'áá/ 'just, only' and /-tʰàà/ 'father'. Contour tones are generally found in languages which have a larger number of tone contrasts. When the language makes only two or three tone distinctions, the tones are usually all level.

2 Definition and distribution of the categories

The first distinction made in this chapter is between languages with and languages without tones. For most languages it is easy to determine if the language does or does not make use of tone, but there are surprisingly sharp disagreements in certain cases. For example, Dar Fur Daju (Nilo-Saharan; Sudan) is reported as non-tonal in one source but transcribed with three tone levels in another. Ket (Yeniseian; northern Siberia) is described as having zero, two, four, or eight tones by different authors (there are some differences in the dialects being described, but this does not account for the differences of opinion on the tonal status of the language). Both these languages have been counted as non-tonal in the present chapter since the opinion that they lack tones seems to be the most well supported (see Thelwall 1981 and Feev 1998 respectively). Other languages have clear word-level pitch phenomena but with limited function, or with roles that look more like stress in that they highlight a particular syllable of a word. Norwegian, Japanese, Ainu, and Oneida (Iroquoian; New York State) are among languages of this kind. These languages are classified here as tonal, but are perhaps only marginally so.

Of the 526 languages included in the data used for this chapter, 306 (58.2 per cent) are classified as non-tonal. This probably underrepresents the proportion of the world's languages which are tonal since the sample is not proportional to the density of languages in different areas. For example, from the large Niger-Congo family of Africa there are 68 languages in the sample, 5 of which are non-tonal (Swahili, Diola-Fogny, Koromfe, Wolof, and Bisa) and the remainder tonal. The *Ethnologue* (Grimes 2000) lists 1489 Niger-Congo languages, so less than 5 per cent of the Niger-Congo languages are included. Of the Indo-European languages of western and central Europe, 16 are included (5 Romance, 3 Germanic, 3 Slavic, 2 Celtic, 1 Baltic, Greek, and Albanian). In these Indo-European groups the *Ethnologue* lists a total of 145 languages (7 Celtic, 58 Germanic, 48 Italic, 18 Slavic, 7 Greek, 4 Albanian, and 3 Baltic languages), so that over 10 per cent of the Western European languages listed are included, only two of which are tonal or marginally so and the rest non-tonal. If, correspondingly, 10 per cent of the Niger-Congo family had been included, 80 additional tone languages would have been included.

1. No tones	306
2. Simple tone system	132
3. Complex tone system	88
total	526

Languages without tones predominate in the western part of the Eurasian landmass, including South Asia, in the more southerly regions of South America, and in the coastal area of northwestern North America. In this last area great genealogical diversity exists among the indigenous languages, but tone is almost entirely absent. In addition, no Australian language has been reported to be tonal.

The languages with tones are divided into those with a simple tone system—essentially those with only a two-way basic contrast, usually between high and low levels—and those with a more complex set of contrasts. About a quarter of the languages (132, or 25.1 per cent) have simple tone systems. This includes 12 languages which appear to meet the definition of being tonal only marginally. With better information a few of these might end up being classed as non-tonal. Less than a fifth (88, or 16.7 per cent) have complex tone systems. Tone languages have marked regional distributions. Virtually all the languages in Africa are tonal, with the greater number having only simple tone systems, although more complex systems are not unusual, especially in West Africa. Languages with complex tone systems dominate in an area of East and South-East Asia. Several clusters of languages with tones occur in South, Central, and North America. A number of the languages of New Guinea are also tonal, or at least marginally so.

3 Relationships to other phonological properties

It is interesting to examine whether there is a relationship between tonal complexity and other measures of phonological complexity. Comparisons were made with the number of consonants, the number of vowel qualities, and the complexity of syllable structures, as described in Chapters 1, 2, and 12 respectively. The number of languages involved in these comparisons is 522 for consonant inventories and 523 for vowel quality inventories.

The mean number of consonants rises with increasing tonal complexity, as shown in Table 1. A simple test of the statistical difference between these means shows that only the difference between

languages with complex tone systems and languages with no tones is significant. Two geographical areas show a relatively high coincidence of larger consonant inventories and tone, Africa and East and South-East Asia.

Table 1 Mean number of consonants by tone system

Tone system	Number of consonants
Complex (n = 88)	26.0
Simple (n = 132)	23.3
None (n = 302)	22.1

There is also a correlation between increasing degrees of elaboration of the tone system and an increase in the mean number of vowel qualities in the languages sampled (all pairwise comparisons are highly significant using a simple comparison of means). Table 2 shows these results. As noted in Chapter 2, Africa and the eastern part of Asia are also areas where larger than average vowel inventories dominate. On a smaller geographical scale, within New Guinea the presence of tone and larger vowel inventories are also correlated. Australia is a stronghold of small vowel systems and non-tonality. In the Americas and the remaining areas of Eurasia the relationship between tone and number of vowel qualities is unsystematic.

Table 2 Mean number of vowel qualities by tone system

Tone system	Number of vowel qualities
Complex (n = 87)	7.05
Simple (n = 130)	6.28
None (n = 306)	5.58

Tables 1 and 2 indicate that tonal complexity and complexity of segment inventory tend to go hand in hand across the set of languages surveyed. In contrast, elaboration of tonal contrasts tends to be inversely related to complexity of syllable structure as defined in Chapter 12. There are 471 languages in the survey for which data on both tone system and syllable structure is present. The results are shown in Table 3, which gives the raw numbers of languages which combine a particular pairing of tone and syllable categories. The percentage of the row total is also shown for each cell, so that the proportion of languages belonging to each tone category which fall into each of the syllable complexity categories can be readily appreciated. For each tone class moderate syllable structure is most frequent, as the middle column shows, but as tonal complexity decreases the proportion of languages with moderate syllable complexity declines going down the column, from just over three quarters to just under a half. On the other hand, the proportion of languages with complex syllable structure increases as tone complexity decreases. Thus it is clear that complex tone systems are strongly correlated with the occurrence of moderate rather than complex syllable structure, whereas non-tonal languages are considerably more likely to have complex syllable structure; languages with simple tone systems fall in between. Tone category does not, however, show any consistent relationship to the occurrence of simple syllable structure, but there are rather few languages concerned.

Table 3 Contingency table of tone system categories (down) and syllable complexity categories (across)

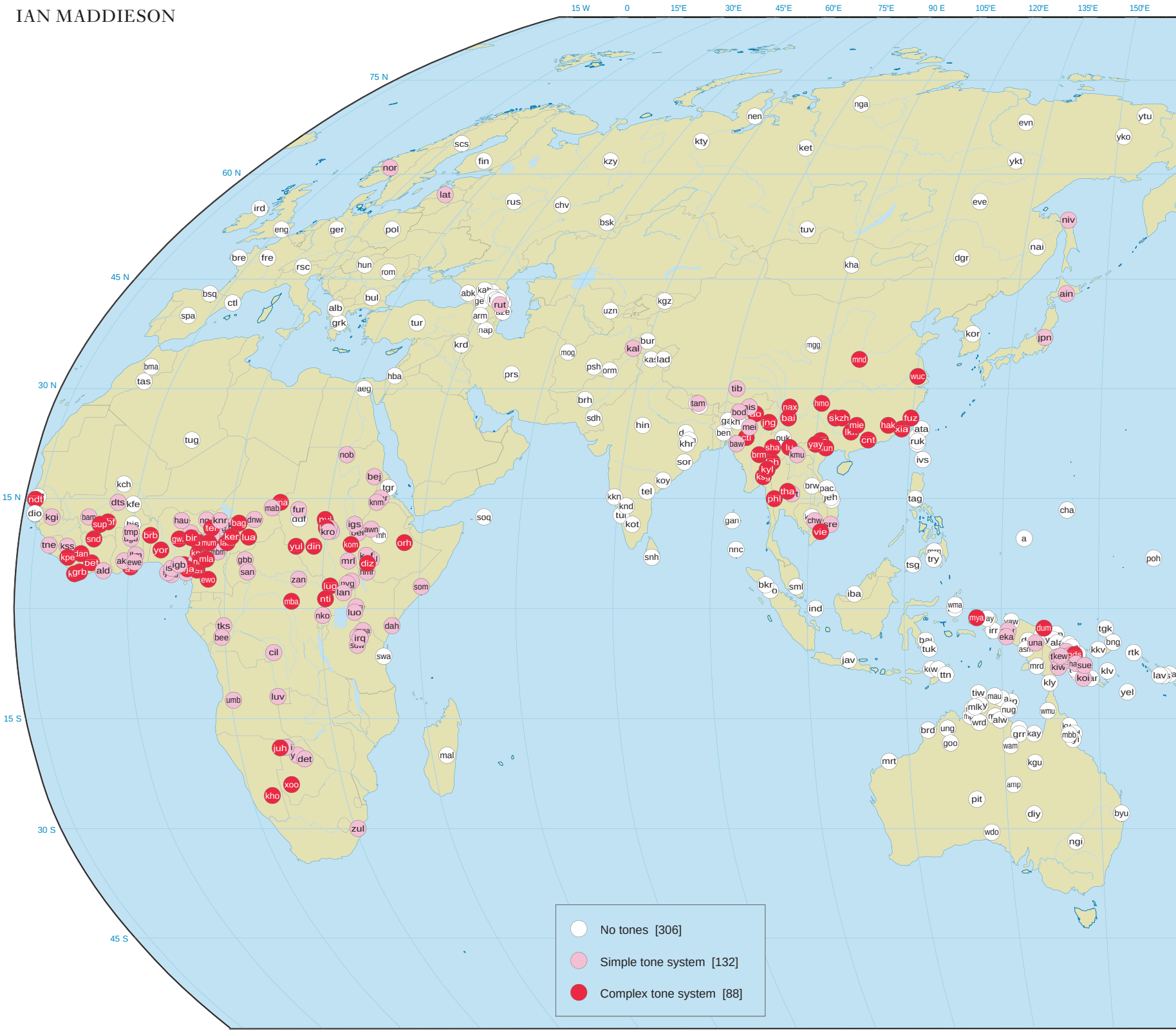
	Syllable			
Tone system	**Complex**	**Moderate**	**Simple**	**Row total**
Complex	8 *(10.4%)*	58 *(75.3%)*	11 *(14.3%)*	77 *(100%)*
Simple/Marginal	23 *(19.3%)*	75 *(63.0%)*	21 *(17.6%)*	119 *(100%)*
None	112 *(40.7%)*	135 *(49.1%)*	28 *(10.2%)*	275 *(100%)*
Column total	143	268	60	471

The pattern shown by Table 3 has a major geographical basis in the high frequency of languages with complex syllable structures in the western part of Eurasia and in the north-west of North America, both areas with few languages having tone. Complex tone systems in Asia are in an area where moderately complex syllable structure dominates.

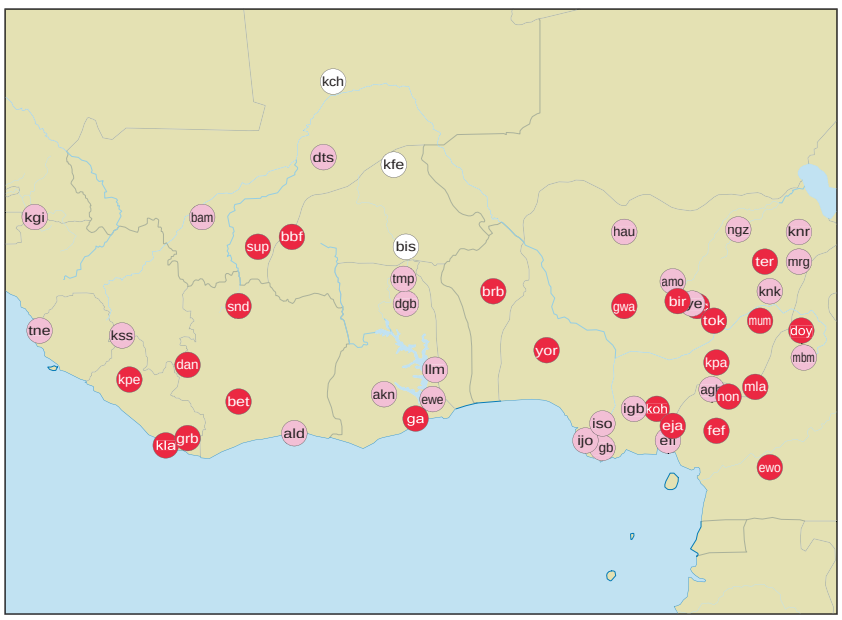
Each of the Tables 1–3 examines a simple overall correlation between the types of tone systems found and another phonological property across all the languages sampled, but it should be borne in mind that this type of analysis does not address the question of whether any of the correlations found are due to fortuitous overlaps of inherited and areally spread properties or represent systematic design features of languages. These results do, however, suggest that tonal complexity might be differently related to segment inventory complexity than it is to syllabic complexity, and that further analysis of these patterns would be worthwhile.

13 Tone

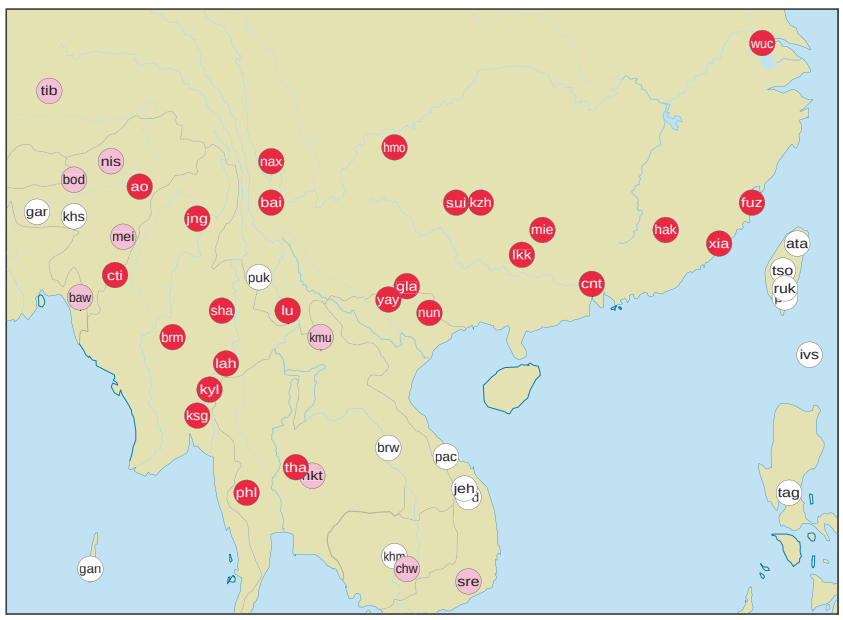
IAN MADDIESON

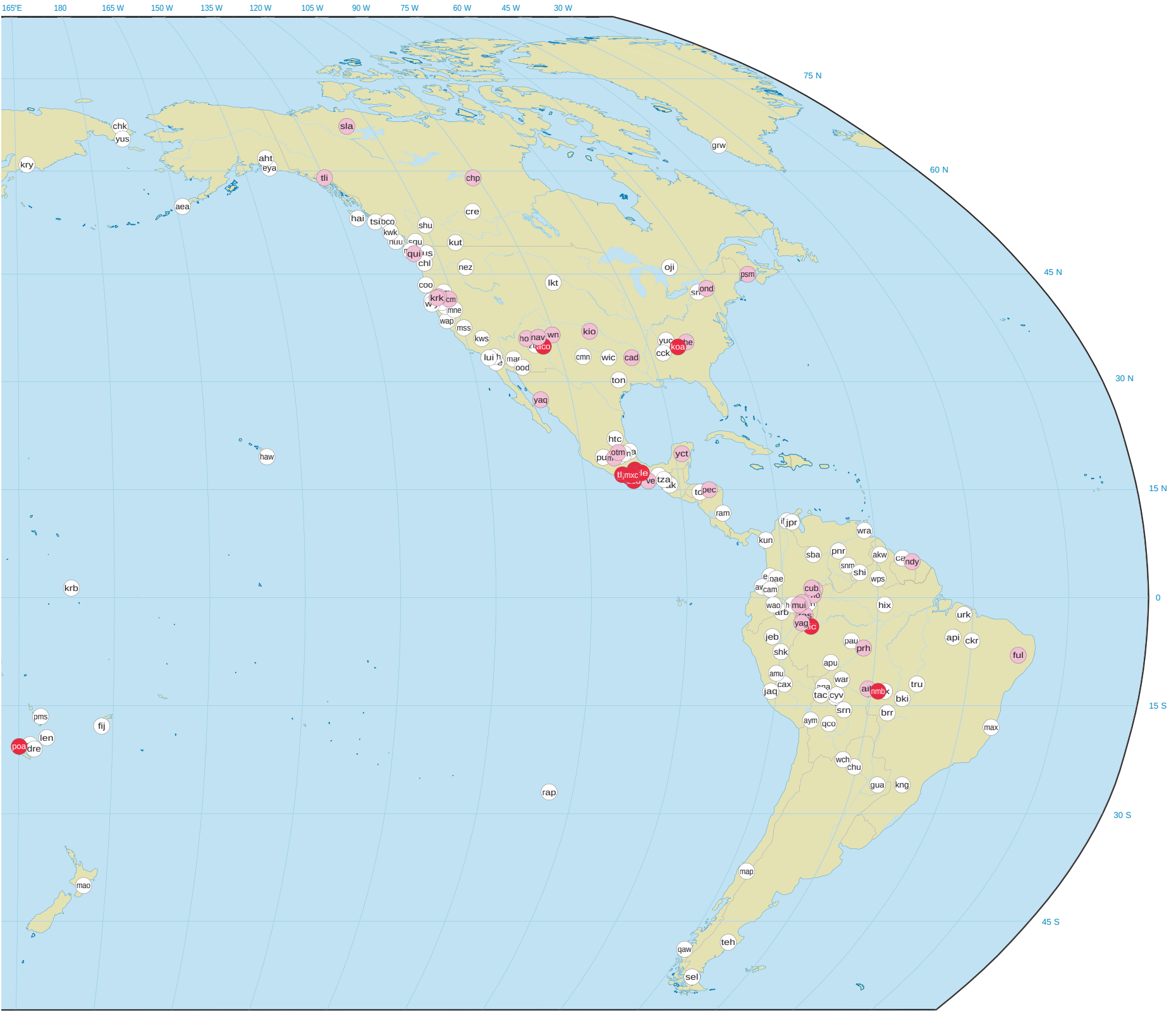


West Africa



mainland South-East Asia





New Guinea and northern Australia

Mesoamerica

14 Fixed Stress Locations

ROB GOEDEMANS AND HARRY VAN DER HULST

1 Stress

This and the next three chapters report on **word stress** and **word rhythm** in the world's languages. The data were collected for a project called **StressTyp** and are also available on-line (see Goedemans and van der Hulst, forthcoming). The StressTyp database contains 510 languages, based on written primary and secondary sources (such as grammars and typological studies). StressTyp naturally only documents languages that have stress and therefore does not provide data about the numerous languages that are reported to lack stress, often tonal languages. In these chapters, ancient or extinct languages are excluded in accordance with the general policy of this atlas, so that the maps are based on a sample of 500 languages.

The term **stress** is used here to refer to an abstract property of syllables within the domain of "words" (cf. Dixon and Aikhenvald 2002 for discussions of the notion *word*). A stressed syllable is likely to be pronounced with more prominence than unstressed syllables. Prominence may involve greater loudness, higher pitch, greater duration, or greater accuracy of articulation (most notably in vowels). Typically, words have at most one syllable that stands out in this way over all other syllables. This syllable is said to have **primary stress** (or **main stress**). Chapters 14–16 deal with the location of the syllable that bears primary stress in the word. In addition, provided that words are long enough, many languages show so-called **nonprimary** or **secondary stresses** (with various rhythm types) on additional syllables. **Rhythm types** are discussed in Chapter 17.

There are, as might be expected, many ways to group stress systems into types. The choices that we have made here represent one (we hope, reasonable) typology, but one that involves lumping together types that one might want to keep separate, or classify differently, in a more elaborate study (cf. §4.2 in this chapter). For reasons of visual clarity, however, we had to limit the number of categories.

We also need to emphasize that we have not considered the role of morphological structure, which, as is well known, may codetermine the location of stress. Our data on the languages in StressTyp are mostly based on stress locations in simplex words or complex words whose morphological structure is not relevant to stress (according to our sources). In some languages, like English, words with certain classes of affixes behave just like simplex words as far as stress is concerned. In other languages, all affixation is like this. We hasten to add that the interplay between stress and morphological structure is in need of more research and documentation (cf. Revithiadou 1999).

2 Defining the values

Somewhat more than half of the languages in our sample have **fixed stress**. By this we mean that stress is located on the same syllable in each word. The location is independent of the **weight** (usually **quantity**) of the syllables in the word and is determined with reference to a word edge only. (Languages with quantity-sensitive stress rules are discussed in Chapter 15.) Among the languages with fixed stress, we distinguish six different types. Thus, Map 14 shows the seven values given in the feature-value box:

1. No fixed stress (mostly weight-sensitive stress)	219
2. Initial: stress is on the first syllable	92
3. Second: stress is on the second syllable	16
4. Third: stress is on the third syllable	1
5. Antepenultimate: stress is on the antepenultimate (third from the right) syllable	12
6. Penultimate: stress is on the penultimate (second from the right) syllable	110
7. Ultimate: stress is on the ultimate (last) syllable	50
total	500

A fairly large number of languages have **initial stress**. An example is Cahuilla (Uto-Aztecan; California): *ˈñaʔaˌčeh* 'sit down', *ˈneñukum* 'female cousins'.

A few languages have stress on the **second syllable**. An example is Mapudungun (Araucanian; Chile and Argentina): *t̪iˈpanto* 'year', *eˈlumuˌyu* 'give us'.

Only one language in our sample, Winnebago (Siouan; Illinois), exhibits stress on the **third syllable:** *hochiˈchinik* 'boy', *waghiˈghi* 'ball' (see also Hayes 1995).

An example of a language with **antepenultimate stress** is Paumarí (Arauan; Amazonas, Brazil): *raˈbodiki* 'wide', *oniˈmanari* 'seagull'.

Penultimate stress is again fairly common. A language with this pattern is Djingili (West Barkly; Northern Territory, Australia): *biˈaŋga* 'later', *ˌŋuruˈala* 'we all'.

A language that exhibits **ultimate stress** is Weri (Trans-New Guinea; Morobe, Papua New Guinea): *uˌluaˈmit* 'mist', *ˌakuˌneteˈpal* 'times'.

3 Geographical distribution

We observe the following broad generalizations:

(i) European systems (if not weight-sensitive) mostly have initial stress. The percentage of weight-sensitive languages is greater in the south.

(ii) Austronesian systems mostly have penultimate stress.

(iii) Arabic dialects are commonly weight-sensitive.

(iv) South American languages are mostly weight-insensitive, with stress oriented towards the right edge of the word.

(v) North American systems form a rather heterogeneous group.

(vi) Except for the northern part (where we find the non-Pama-Nyungan languages with penultimate stress), Australian systems mainly have initial stress.

4 Theoretical interpretation

4.1 Stress parameters. Several factors determine the location of primary stress:

(1) a. size of the **stress window** (*bounded*: two peripheral syllables; or *unbounded*: the whole word)

b. the location of the window if it is **bounded** (*left* or *right* edge of the word)

c. **nonperipherality** or **extrametricality** (yes/no; if yes: exclusion of a peripheral syllable of the word in determination of the stress window)

d. the **location of stress** within the window (*left* or *right*)

e. syllable weight (yes/no; if yes: *heavy* and *light* syllables)

The role of these "parameters" will be explained in the texts that accompany this and the following maps. We deliberately choose a quasi-formal, parametric approach so that interested readers will come prepared when consulting the StressTyp database on which the present maps are based.

Map 14 deals with languages in which the stress location is fixed. In this study, this means that the window is **bounded**. Stress is located within a two-syllable window that is either on the left side or the right side of the word. Here we do not reckon with syllable weight; we leave that for Chapter 15. In some cases, it appears that the stress window is trisyllabic because stress can occur on the antepenultimate or third syllable (either regularly or, if stress is not fixed, as an option). To "reach" the third position in, students of stress suggest that the rightmost or leftmost syllable in the word can

be disregarded in the determination of the stress window. We refer to this parameter as **nonperipherality** (the term *extrametricality* is also often used). Thus, for languages in which the size parameter = bounded, and the weight parameter = no weight, we can now distinguish the six cases in Table 1 (the symbol σ stands for a syllable, boldface and underlining means stress, parentheses indicate the stress window, and brackets show the word edge):

Table 1 Fixed stress types

Pattern	Window	Nonperiph.	Location	
Initial	left	no	left	[(σ́ σ) σ
Second	left	no	right	[(σ σ́) σ
Third	left	yes	right	[σ (σ **σ́**)
Ultimate	right	no	right	σ (σ σ́)]
Penultimate	right	no	left	σ (σ́ σ)]
Antepenultimate	right	yes	left	(σ́ σ) σ]

Henceforth we refer to a window having a left-location for stress inside the window as **left-headed**. The opposite choice will be called **right-headed**.

Most languages with fixed stress have or tolerate exceptions, which, interestingly, represent small deviations from the fixed rule (here understood as a set of parameter settings). Malagasy, for example, has penultimate stress, but allows words with ultimate stress (marked "location: right-headed") or with antepenultimate stress (marked "nonperipherality: yes"). These alternatives are exemplified in (2).

(2) Malagasy (Austronesian; Madagascar)
ta'nana 'village' *o'me* 'to give' *'satroka* 'hat'

Exceptions with regard to the window parameter are rare, if not absent. For example, it is highly unusual for a language with penultimate stress to have exceptions with initial stress.

This being said, fixed antepenultimate stress is rare. We have only twelve cases in the sample. This does not mean that antepenultimate stress location *as such* is rare in languages of the world. A typical pattern in the Romance languages (except French) is that stress can be ultimate, penultimate, or antepenultimate. Such cases can have a **dominant pattern** (let us say penultimate), with ultimate and antepenultimate being less frequent. If the number of exceptions is so high that the dominant pattern may be obscured, we treat the stress location as **variable** rather than fixed (cf. Chapter 15). The choice between fixed (with exceptions) and variable (with remnants of a dominant pattern) cannot always be made objectively.

The most frequent fixed pattern, then, is penultimate, followed by initial. Both correspond to a left-headed, bounded domain (**σ́**σ), which is a **trochaic** (i.e. "strong-weak") pattern. Antepenultimate stress also involves this pattern, combined with right-edge nonperipherality. It is interesting to note that closely related languages with different fixed stress patterns may share this trochaic pattern, while differing in terms of other parameters. Consider the following three Slavic languages:

Table 2 Fixed stress in three Slavic languages

	Window	Nonperiph.	Location	Pattern
Czech: Initial	left	no	left	[(σ́ σ) σ
Polish: Penultimate	right	no	left	σ (**σ́** σ)]
Macedonian: Antepen.	right	yes	left	(**σ́** σ) σ]

The preponderance of the trochaic type supports the idea that stress in languages with bounded domains is grounded in the same principles that govern the rhythmic alternation of secondary stresses, where, as we will see in Chapter 17, left-headed (or trochaic) patterns are also dominant.

The third most frequent type is ultimate, followed at a distance by second-syllable stress. Both patterns involve a right-headed, **iambic** window (σ**σ́**). Fixed second-syllable stress is very rare, much rarer than ultimate stress. Fixed third-syllable stress is found in only one language, Winnebago.

4.2. Ambiguities. We now wish to draw attention to the fact that the interpretation of stress patterns as being left-headed (within the stress window) or right-headed may not be as straightforward as it seems. There has been a certain amount of discussion concerning the need for right-headed (iambic) patterns, which at first sight seem inevitable for the ultimate and second-syllable stress types. It *could* be said, however, that second-syllable stress results from nonperipherality combined with a left-headed stress domain:

(3) [σ (**σ́** σ) . . .
Siroi (Trans-New Guinea; Papua New Guinea)
am'bonga 'at first'

Since the penultimate type is more frequent than the antepenultimate type, we consider nonperipherality to be a marked phenomenon. Thus, the advantage of the analysis in (3) is that second-syllable stress would fall in the same category as antepenultimate stress, both being relatively rare. But this "explanation" for the rarity of the second-syllable pattern presupposes that we eliminate iambic, primary-stress domains altogether, which, in turn, raises the question of how we will deal with the rather frequent ultimate type.

Here is a possible way of re-analysing these ultimate cases. As we have specified above, languages may choose an unbounded domain for stress assignment, in which case stress may fall on any syllable in the word, provided that this syllable is heavy, while (preceding or following it, depending on the language) others are light. As a result, stress may end up in a location that is clearly outside the reach of a bounded window, as in Northern Haida, where stress falls on the rightmost heavy syllable (with high tone) in the word, or the last if there are no heavy syllables.

(4) [σ σ σ σ́$_h$ σ σ σ]
Northern Haida (Na-Dene; British Columbia, Canada)
kil,ɬaldaanh'láng,id?wa,gan
'caused.POT.be.slaves.INCEP'

Indeed, the usual idea is that unbounded systems are all weight-sensitive. However, it *might* be argued that the choice of an unbounded window may also coincide with the choice of not being sensitive to syllable weight, in which case a language with fixed ultimate stress could be characterized as being unbounded, (5a), rather than being bounded, (5b):

(5) a. [σ σ σ σ σ́] b. [σ σ σ (σ σ́)]

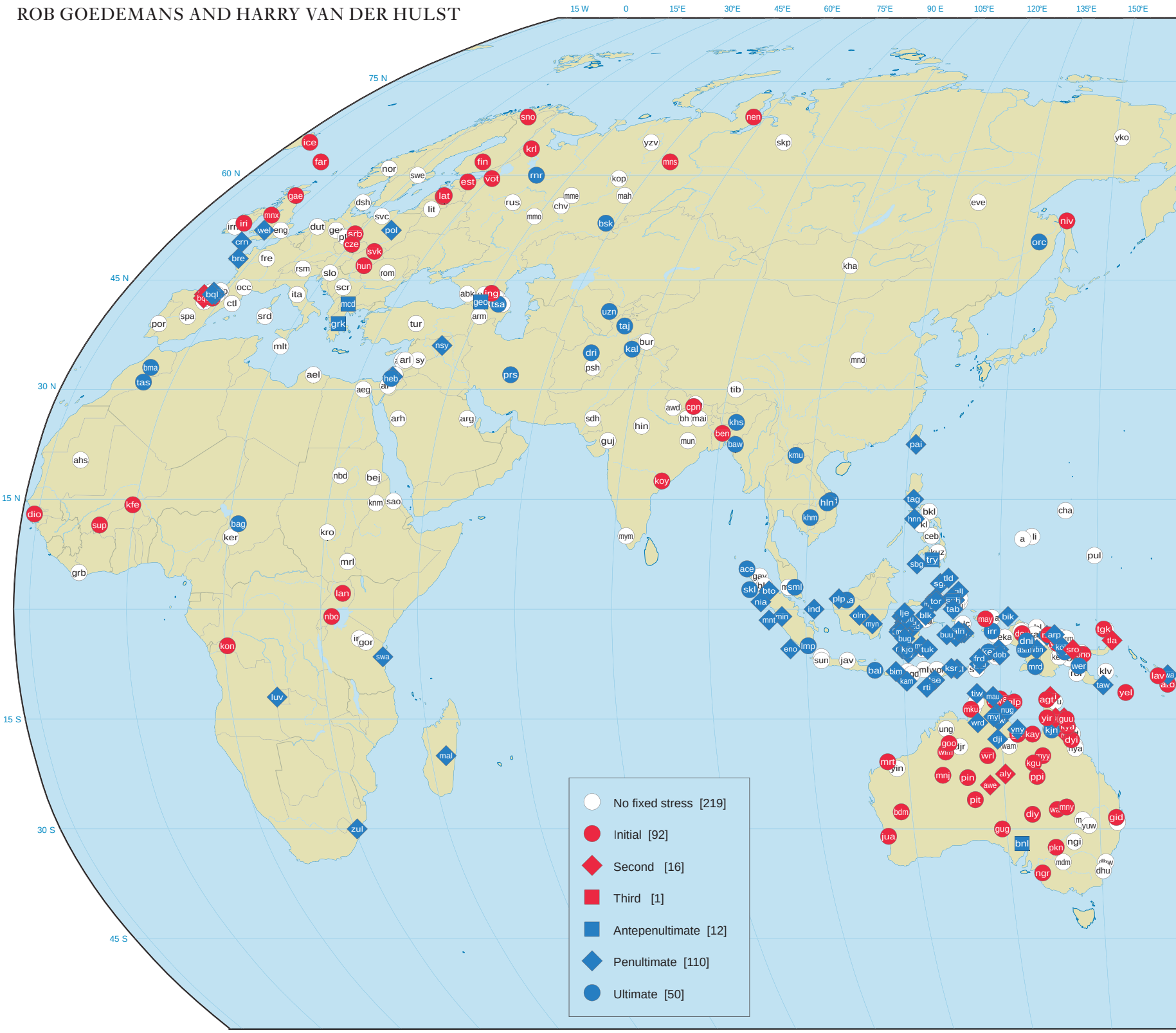
Of course, once this discussion is opened, one might wish to re-examine penultimate stress or second-syllable stress cases in order to determine whether these perhaps combine unboundedness of the stress window with nonperipherality:

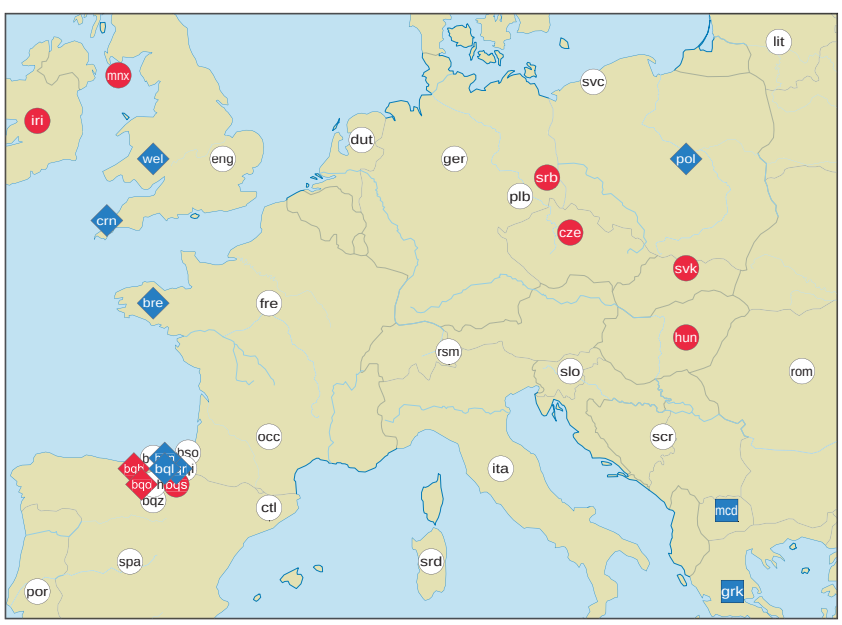
(6) a. [(σ σ σ **σ́**) σ] b. [σ (**σ́** σ σ σ)]

Although the discussion has now assumed a rather theoretical and hypothetical flavour, unfit for a data-oriented survey, we believe that it is important to see that classifications based on surface patterns can be dangerous and misleading, especially if one assumes a direct mapping from descriptive categories onto theoretical categories. We add, however, that the descriptive categories used for the maps are not affected by these different theoretical interpretations.

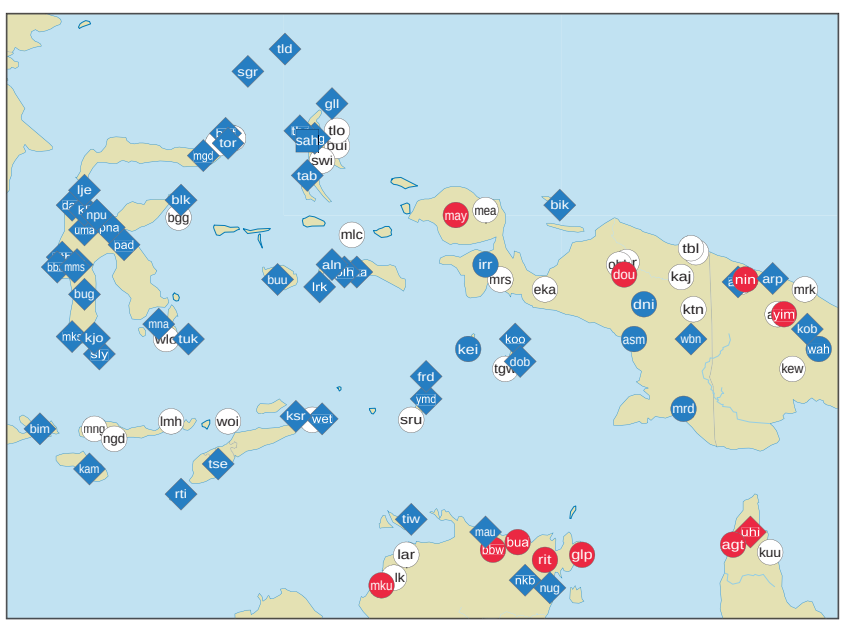
14 Fixed Stress Locations

ROB GOEDEMANS AND HARRY VAN DER HULST

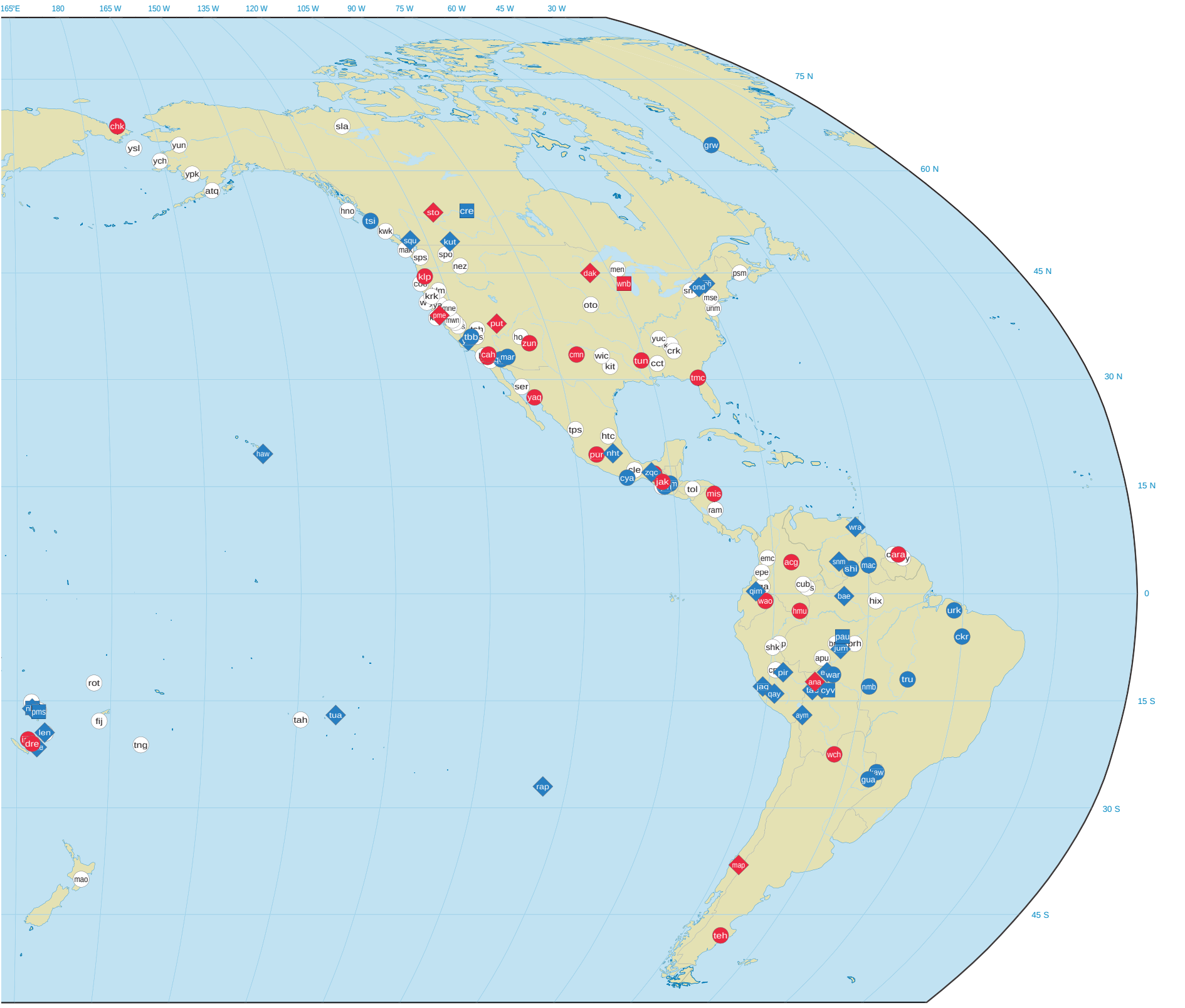


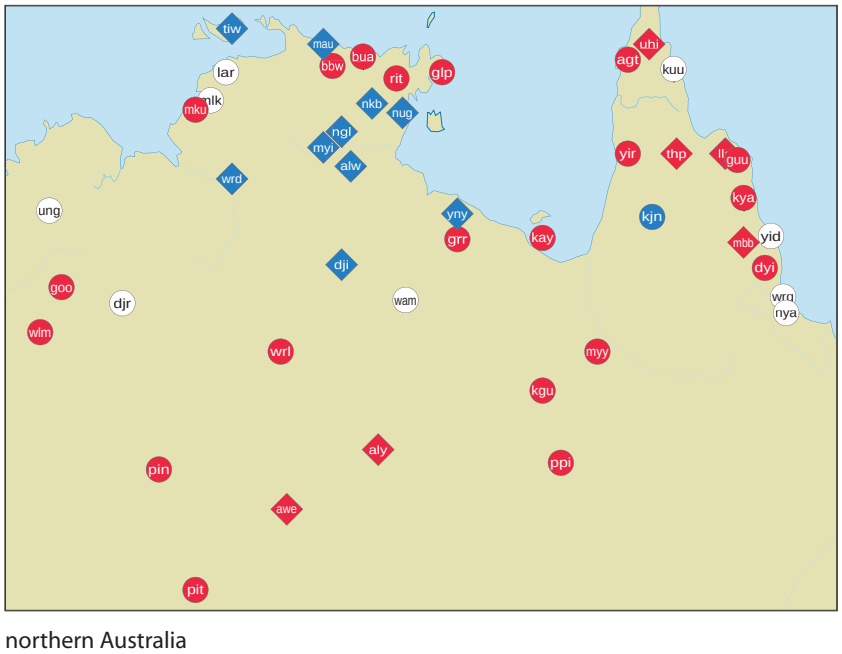


Europe

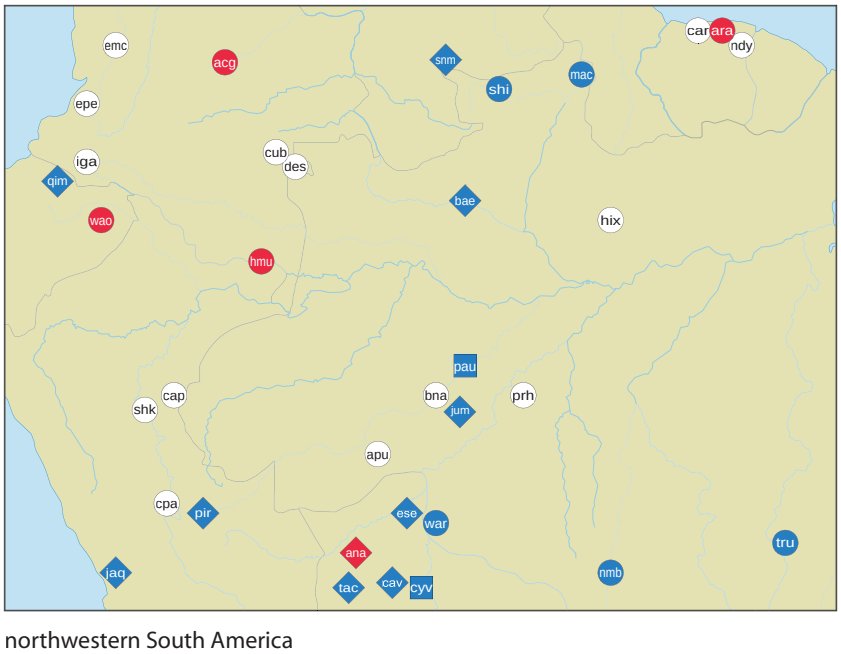


eastern Indonesia





northern Australia



northwestern South America

15 Weight-Sensitive Stress

ROB GOEDEMANS AND HARRY VAN DER HULST

1 Syllable weight

In this chapter we focus on systems in which the placement of stress is sensitive to properties of syllables. In general, if a stress system is sensitive to such properties, this means that syllables fall into two classes, called **heavy** and **light** syllables. Light syllables are prototypically CV; heavy syllables are prototypically either CV: or CVC or both, depending on the language (see Chapter 16). Heavy syllables are such that they "attract stress", while light syllables only get stress if they happen to be in the right location in the string of syllables. In some languages, syllable weight has been argued to form a scale with more than two values.

We can distinguish between various types of stress-attracting properties, such as **intrinsic properties** involving vowel length, syllable closure, or "prominence", and **extrinsic properties** like rhythmic strength and "diacritic weight" (or "accents"). In this chapter all these factors are lumped together. Chapter 16 offers a map that distinguishes the various kinds of syllable weight.

2 Defining the values

For Map 15 we distinguish eight types:

1.	Left-edge: stress is on the first or second syllable	37
2.	Left-oriented: the third syllable is involved	2
3.	Right-edge: stress on ultimate or penultimate syllable	65
4.	Right-oriented: the antepenultimate is involved	27
5.	Unbounded: stress can be anywhere in the word	54
6.	Combined: both right-edge *and* unbounded	8
7.	Not predictable	26
8.	Fixed stress (no weight-sensitivity)	281
	total	500

2.1 Bounded systems (types one through four). Let us first consider cases in which the weight-parameter is positive, while the window is bounded (see Chapter 14, §4.1, on stress parameters). This gives us weight-sensitive, bounded systems. With stress being sensitive to a distinction between light and heavy syllables, four logical possibilities arise for each stress window. We exemplify this with reference to a right-edge, bounded window (H and L stand for heavy and light syllable, respectively; boldface type means that stress falls on the given syllable):

(1) a. (**H** L)] b. (L **H**)] c. (H H)] d. (L L)]

The universal property of a weight-sensitive system is that in cases (1a) and (1b) stress will always be located on the heavy syllable. We immediately see why weight-sensitive systems are crucially different from the systems presented on Map 14: there is no fixed stress location in every word, but rather, the location depends on the presence and location of heavy syllables within the stress window. For each given word the location is fully determined, but for the lexicon as a whole we see that various stress locations occur, given that words have different arrangements of heavy and light syllables. Languages with right-edge windows and weight-sensitivity may differ from each other with respect to (1c) and (1d). In case (1c) the stress will usually fall on the final (i.e. rightmost) heavy syllable so that, in general, it can be said that stress falls on the rightmost heavy syllable in the stress window. In (1d), the most usual case is to have stress on the left-hand light syllable (i.e. the familiar trochaic pattern). Overall then, the most common right-edge, weight-sensitive system would be described in (2i). However, the other logical options for (1c) and (1d) are also possible, so that we get a total of four possible right-edge systems:

(2) i. a. (**H** L)] b. (L **H**)] c. (H **H**)] d. (**L** L)] Epena Pedee [38]
ii. a. (**H** L)] b. (L **H**)] c. (H **H**)] d. (L **L**)] Yapese [5]
iii. a. (**H** L)] b. (L **H**)] c. (**H** H)] d. (**L** L)] Awadhi [8]
iv. a. (**H** L)] b. (L **H**)] c. (**H** H)] d. (L **L**)] Aklanon [4]

(The frequencies of these patterns in StressTyp are given between square brackets.)

The differences between the four types result, then, from different parameter settings for the two possible cases in which the window contains syllables of equal weight:

(3)	(H H)	(L L)	
i.	right-headed	left-headed	(Right/Left)
ii.	right-headed	right-headed	(Right/Right)
iii.	left-headed	left-headed	(Left/Left)
iv.	left-headed	right-headed	(Left/Right)

Here are some illustrative examples for the most common pattern (i.e. (2i)), from Epena Pedee (Choco; Colombia):

(4) Epena Pedee (long vowels form heavy syllables (H))

'warra	'son'	*war'raa*	'flavourful'
'taama	'snake'	*tee'soo*	'long'

Interestingly, the same parameter-setting options create four possibilities at the left-edge window:

(5) i. a. [(**H** L) b. [(L **H**) c. [(**H** H) d. [(L **L**) Ossetic [9]
ii. a. [(**H** L) b. [(L **H**) c. [(**H** H) d. [(**L** L) Malayalam [15]
iii. a. [(**H** L) b. [(L **H**) c. [(H **H**) d. [(L **L**) Archi [1]
iv. a. [(**H** L) b. [(L **H**) c. [(H **H**) d. [(**L** L) Capanahua [3]

The following examples from Malayalam (Dravidian; southern India) illustrate the most common pattern (5ii):

(6) Malayalam (long vowels make syllables heavy)

'kutira	'horse'	*paʈ'ʈaaʟak,kaaran*	'soldier'
'kuuʈʈam	'crowd'	*'aakaaçam*	'sky'

In both Epena Pedee and Malayalam stress falls on the most peripheral heavy syllable, with the trochaic pattern being the choice if no heavy syllable is available within the window.

StressTyp differentiates between all these possibilities. However, on Map 15, all cases in which stress varies (depending on weight) between two locations within a window are grouped under the labels **left-edge** and **right-edge**. In other words, for each edge we have collapsed four logical possibilities.

Next, we turn to the categories **left-oriented** and **right-oriented**, which differ from left-edge and right-edge in the following way. We use the term *oriented* when the possible stress locations include "the third syllable in", i.e. the antepenultimate syllable for right-oriented systems, and the third syllable for left-oriented systems. Inclusion of these syllables occurs in several types of cases.

First, the class of right-oriented languages includes cases in which nonperipherality is involved, as in classical Latin:

(7) Classical Latin

a. (**H** L) σ]	b. (L **H**) σ]	c. (H **H**) σ]	d. (**L** L) σ]
do'mesticus	*re'feːcit*	*per'fectum*	*'anima*

Secondly, also included in left-/right-oriented are languages that, in principle, assign primary stress in a two-syllable window (without extrametricality), but which include exceptional cases that have, for example, antepenultimate stress. Elsewhere, we have labelled such systems "broken window systems". A typical case is a language that has bounded weight-sensitive stress, as described above, but where in the case of H(LL)] stress is assigned to the antepenultimate syllable. The heavy antepenultimate syllable "pulls" stress outside the stress window.

Thirdly, it may happen that stress may occur on all three peripheral syllables dependent on weight factors. A clear case is Pirahã (Mura; Amazonas, Brazil), where stress falls on the strongest of the last three syllables in a word.

Fourthly, we have also included languages in which the location of stress can be ultimate, penultimate, or antepenultimate not because intrinsic weight is relevant, but simply because stress is lexically marked within a peripheral or near-peripheral right-edge domain:

(8) a. σ (σ **σ**)] b. σ (**σ** σ)] c. (**σ** σ) σ]

(We distinguish such a language, which could have a weak preference for one of the locations, such as penultimate, from a language in which stress is clearly penultimate, while some exceptional cases of antepenultimate and ultimate are present; recall Chapter 14, §4.1.) The rationale behind including this type in the bounded weight-sensitive systems is that lexical marking of syllables functions much like intrinsic weight. A lexically marked syllable will attract the stress within a domain, just as a heavy syllable does. We therefore refer to lexical marking as **diacritic weight** (see Chapter 16).

2.2 Unbounded systems (type five). This category, also represented on Map 15, has an unbounded stress window comprising the whole word (with the possible exclusion of a peripheral syllable). These systems are especially interesting because the location of stress is not restricted to syllables that are near the edge of the word. The reason why unbounded systems are represented together with weight-sensitive bounded cases is that, in StressTyp, all unbounded systems are weight-sensitive. In Chapter 14 we discussed the possibility of treating systems that we classified as weight-insensitive bounded systems, as unbounded in certain cases. However, we have not done this in StressTyp. Hence, in StressTyp all unbounded systems are weight-sensitive.

It turns out that the location of stress in unbounded systems follows exactly the same principles that determine stress location in bounded weight-sensitive systems. Again there are four types, as shown in (9) (compare (3)). The only difference is that the domain can now contain more than two heavy syllables (since its size is not limited to two syllables):

(9)

	All H	All L	
i.	right-headed	left-headed	(Right/Left)
ii.	right-headed	right-headed	(Right/Right)
iii.	left-headed	left-headed	(Left/Left)
iv.	left-headed	right-headed	(Left/Right)

A more "user-friendly" description of (9i) would be:

(10) Stress the rightmost heavy syllable in the word, and if there are no heavy syllables, stress the first syllable

All four logical options are attested in the languages of the world:

(11)

i.	Right/Left:	Dongolese Nubian, Huastec, Meadow Mari	[13]
ii.	Right/Right:	Aguacatec, Golin, Hill Mari	[10]
iii.	Left/Left:	Khalkha, Murik	[23]
iv.	Left/Right:	Komi-Permyak, Kwakw'ala	[7]

Here are some examples from Dongolese Nubian:

(12) Dongolese Nubian (both long vowels and non-word-final codas render syllables heavy)

ˈnosogid	'old age'
kɛmiŋˈgaːr	'all four'
aroˈgittɛl	'whitish'

On the map, all four types are collapsed into one category of unbounded systems.

The **combined** category (type six) contains languages in which the stress rule switches from bounded to unbounded (or vice versa) under predictable circumstances—for instance, when the bounded domain contains only light syllables, or in the assignment of secondary stress. The relevance of this category to the theory is high, but the discussion is too elaborate to present here.

In the cases discussed so far, syllable weight is determined by the intrinsic properties of syllables. However, as with bounded systems, we also reckon with lexically marked weight, called "diacritic weight". Consider the stress system of Russian. It is well known that in this language specific syllables in morphemes can be 'marked' for stress (such marks are often called "accents"). If a morphologically complex word ends up with more than one such mark, it will be the first that will receive primary stress. If, on the other hand, no syllable is accented because all morphemes that constitute the word are unaccented, the first syllable of the word is stressed. In other words, Russian is a Left/Left system just like Khalkha. The difference is that in Russian, syllable weight is diacritic, whereas in Khalkha it is phonological (involving vowel length).

A final variable stress type is formed by the so-called **count systems**, in which stress depends on the number of syllables in the word. We find such a system in Malakmalak (Australian, Daly; Northern Territory, Australia), which stresses the first syllable if the number of syllables is even, otherwise the second (e.g. *ˈmunanˌkara* 'beautiful', *anˈkiniˌyaŋka* 'you and I will stand'). We analyse this system by assigning trochaic feet from right to left (leaving one syllable "unfooted" in the odd-syllable case), and placing the main stress on the head of the foot that was assigned last. Thus, "heaviness" is determined by **rhythmic strength**, rather than by intrinsic or diacritic weight. Hence the introduction of a third type of weight: **rhythmic weight**. Both bounded and unbounded analyses of this type are possible (see van der Hulst 1996).

2.3 Nonpredictable stress (type seven). Map 15 covers a further category of stress systems that we have labelled **not predictable**. Here we group together languages that have no predictability in stress location whatsoever (cf. Burushaski, Spokane, Usan) and languages in which words are said to have several stresses (rhythmic or weight-determined), with none of these being primary, or all being equal. The former category would qualify as having **free stress**. Not to be confused with "free" stress would be a system like Russian (see above), which, while depending on lexically marked, diacritic weight, is not free; only the lexical marks (accents) per *morpheme* are said to be free. Cases that involve no primary stress, or equal stresses, are not uncommon. We find this for instance in St Lawrence Island Yupik (Eskimo; Alaska) and in Hixkaryana (Carib; Brazil):

(13) St Lawrence Island Yupik
ˈaːŋqaχˈɬaχɬaŋˈyuxtuq 'he wants to make a big ball'

We suspect that in such cases additional study might still reveal the primacy of one of the stresses. The same may hold true for the free stress systems. Sometimes one location in a "no main stress" language is always stressed, while other locations, for instance, are stressed only when there is a long vowel, as in Gidabal (Pama-Nyungan; Queensland, Australia). In such cases StressTyp already treats the recurring stress as a primary one.

3 Geographical distribution

A number of general tendencies can be observed:

(i) The overall number of unbounded systems is higher than one might suspect and they occur all over the world (except in Austronesian languages).

(ii) Austronesian languages show a clear preference for right-oriented stress, which, added to their preference for penultimate fixed stress (Chapter 14), makes the right edge the clear preference all around.

(iii) The northern part of Australia (with preference for fixed penultimate stress) lacks right-oriented cases.

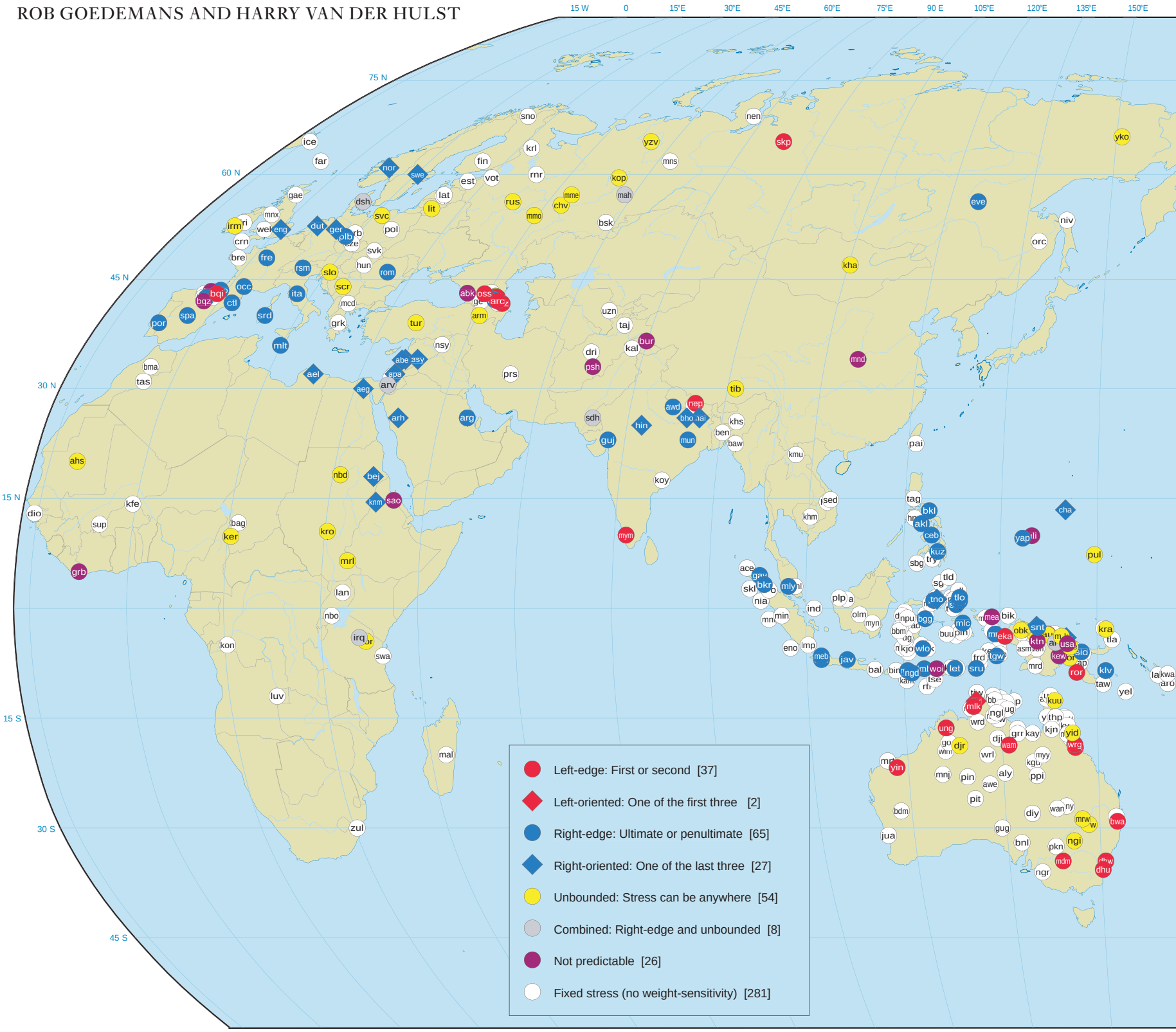
(iv) Australian languages as a whole show very little weight-sensitivity.

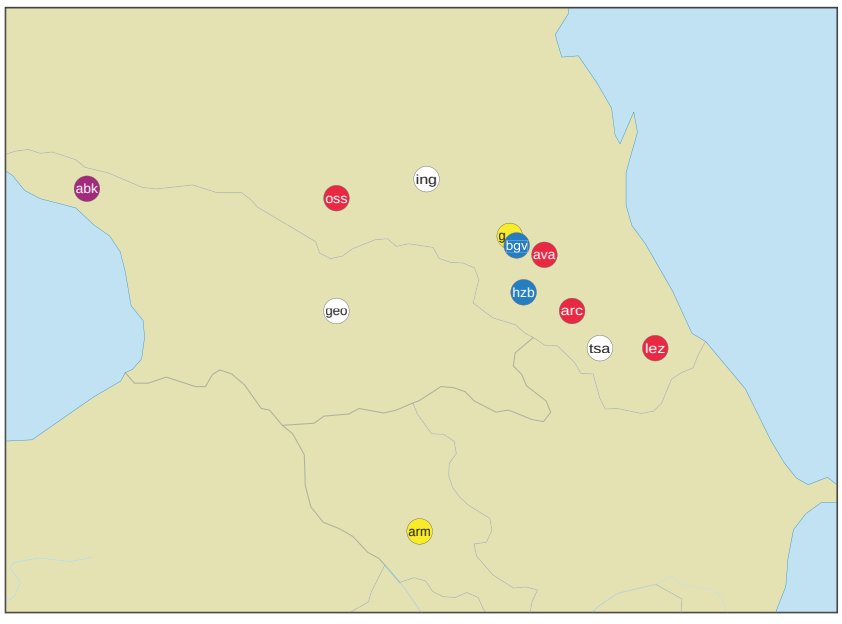
(v) European languages do not have left-oriented systems.

(vi) African languages are only unbounded and right-oriented.

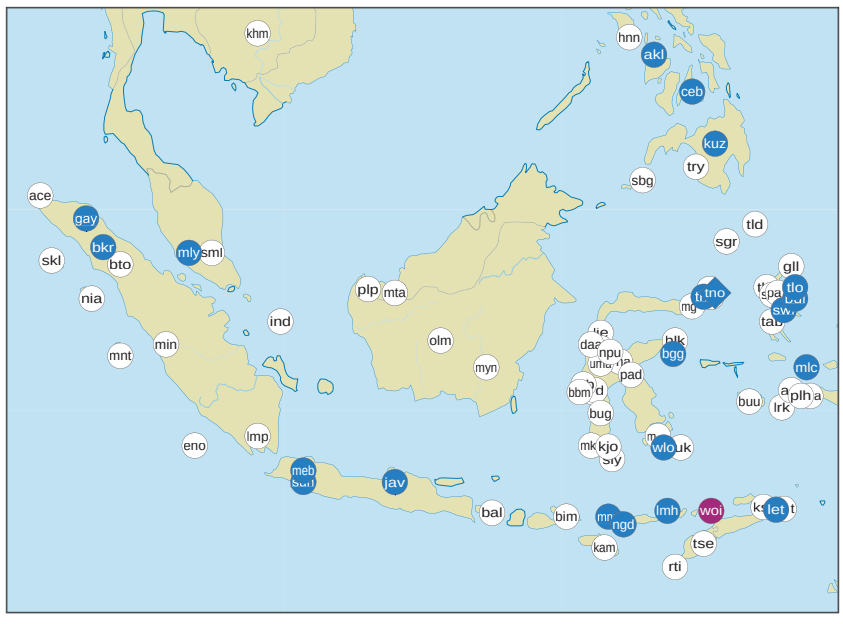
15 Weight-Sensitive Stress

ROB GOEDEMANS AND HARRY VAN DER HULST

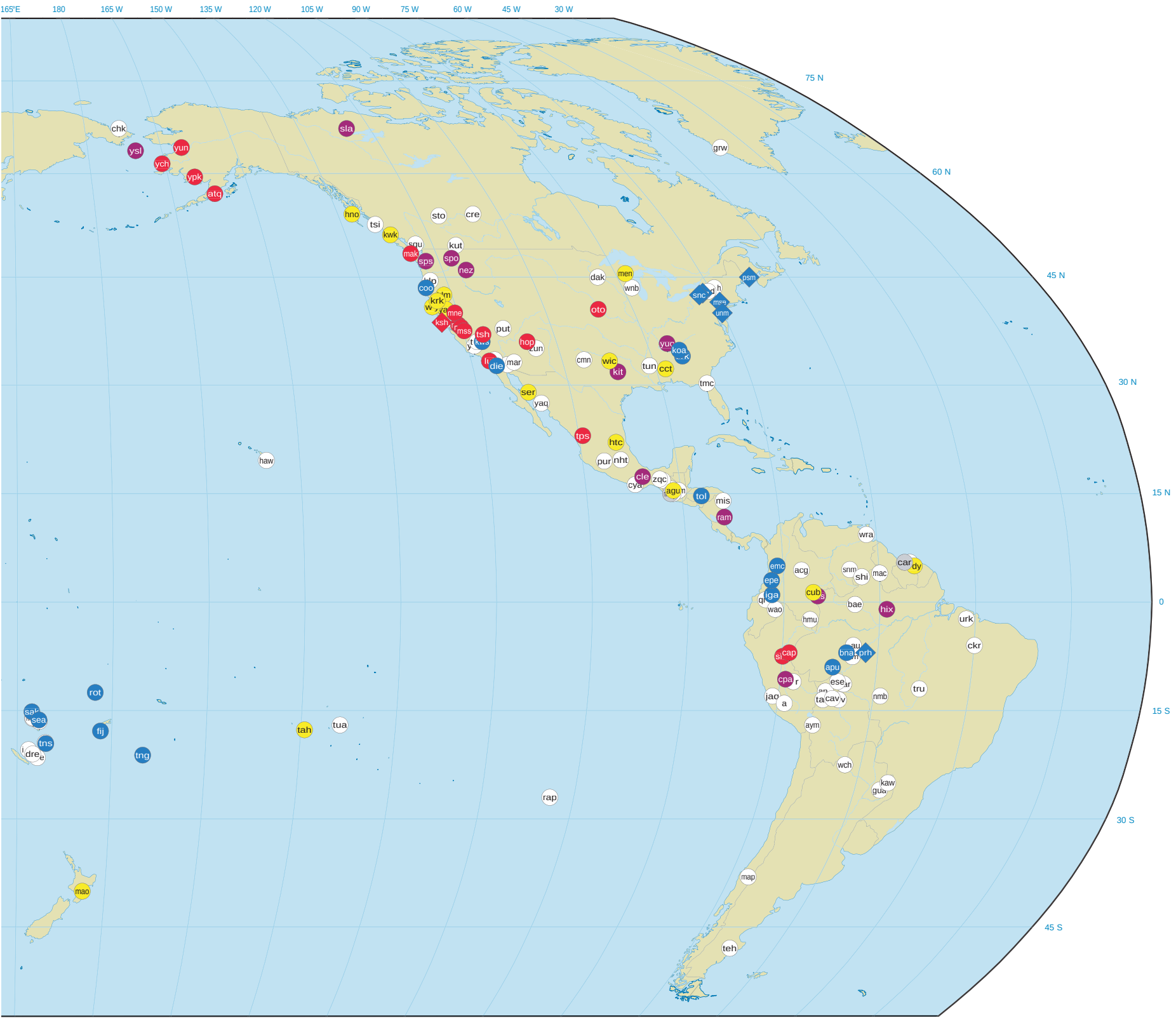




Caucasus



Sumatra, Borneo and surrounding areas



New Guinea

western United States

16 Weight Factors in Weight-Sensitive Stress Systems

ROB GOEDEMANS AND HARRY VAN DER HULST

1 Phonological weight

This map is concerned with the notion of **phonological weight**. In the two preceding chapters, we have made a distinction between several types of weight:

(1) a. intrinsic weight (due to properties of the syllable)
b. rhythmic weight (due to rhythmic footing)
c. diacritic weight (due to lexical marking of 'accents')

Diacritic weight equals lexical marking. Rhythmic weight was assumed for **count systems** (cf. Chapter 15, §2.2). The first-mentioned category, **intrinsic weight**, needs some further comment. The intrinsic properties that determine weight can differ from one language to the next. An important determinant of weight is **vowel length**, closely followed by **syllable closure**. In principle, the two factors are independent, but often both long vowels and closing consonants will cause syllables to be heavy for primary and/or secondary stress in one and the same language, leaving the light category for open, short-voweled syllables. Vowel length and syllable closure are sometimes referred to as **quantitative weight**.

It is generally assumed that syllables are divided into an **onset** (prevocalic part) and a **rhyme** (the rest), and that the latter part has a **nucleus** (vocalic part) and a **coda** (postvocalic consonant). This can be shown by the following representations, which are relevant for the purposes of quantity:

(2) a. O R b. O R c. O R

N N N Cd

V V V V C

The most common types of heavy syllables are (2b) and (2c), which both have two rhyme units, or, to use a more common term, **moras**. Broadly defined, a mora is a unit of weight, which light syllables have one of, and heavy syllables two.

Our "type 5" languages involve another weight factor, called *prominence*, which we discuss in more detail in §4.4.

2 Defining the values

○	1. No weight, or weight factor unknown	261
●	2. Long vowel: long vowels are heavy for stress	65
●	3. Coda consonant: closed syllables are heavy for stress	18
●	4. Long vowel + Coda: long vowels *or* closed syllables	35
●	5. Prominence: other factors are heavy for stress	41
●	6. Lexical: lexical stress, diacritic weight	38
○	7. Combined: two of the above factors determine weight	42
	total	500

2.1 No weight. In about half of the sample languages (261), syllable weight plays no role for word stress. (This is less than the corresponding figure (281) in Chapter 15 because there we looked only at weight-sensitivity for *primary* stress.)

Weight-sensitive languages are subdivided into five different types. The following five values are shown on the map:

2.2 Long vowel. In this type, only a long vowel makes a syllable heavy. An example of this comes from Iraqw (Southern Cushitic; Tanzania): *ba'ʔeeso* 'bushbucks', *'wawitmo* 'king'. Notice how the closed syllable in the second example does not draw stress onto itself.

2.3 Coda consonant. These are mostly languages that have no long vowels. Hence, only a coda consonant can make syllables heavy. An example of such a language is Amele (Madang, Trans-New Guinea; Papua New Guinea): *'nifula* (kind of beetle) *iti'tom* 'righteous'. Amele stresses the first heavy syllable (second example) or the first syllable (first example). Languages that have long vowels but still use only codas for weight are extremely rare or absent (see discussion in §4.2).

2.4 Long vowel + Coda. For languages that do count codas for weight, and which do have long vowels, the default option is that either one can make a syllable heavy by itself. An example of this type is Hopi (Uto-Aztecan; Arizona). If the first syllable is light, stress falls on the second: *ho'nani* 'badger', but if the first is heavy, it is stressed: *'sipmasmi* 'silver bracelet', *'naatɪhota* 'to hurt oneself'.

2.5 Prominence. Prominence systems form a somewhat mixed collection. There are several nonquantitative salient syllabic properties that can form the basis of heavy–light distinctions. One of the most common is the opposition between full and reduced vowels, as in Chuvash (Turkic; Russia), which stresses the last full vowel or else the first reduced one: *sarla'ka* 'widely', *'ĕslĕpĕr* 'we shall work'. To save space here, we count rhythmic weight systems (1b) among the prominence systems. For further discussion of possible theoretical motivations for counting rhythmic beats as prominence factors see Goedemans (1996).

2.6 Lexical marking. Lexical marking yields systems with unpredictable stress by definition. Because stress is lexical, it can be a distinctive property, as we see in these Pashto (Iranian; Pakistan and Afghanistan) examples: *'guta* 'knot', *gu'ta* 'pochard'.

2.7 Combined. This category is a mixed bag of no less than fifteen different combinations of two of the above weight factors. Combinations can occur if the stress rules are sensitive to different heavy-light distinctions in different positions or circumstances. The large size of the group has no typological meaning, since the systems contained in it are very diverse. We represent them as a single group only because they do not fit the other types, and have the "dual nature" in common. Theoretically, these combination systems are very interesting, so in that respect, the large size of the group is quite relevant. A discussion of this, though, is far beyond the scope of this chapter.

3 Geographical distribution

We notice a number of broad generalizations:

(i) We find no lexical stress in South America and Australia.
(ii) Quantitative weight is found in all areas and families.
(iii) We find a relative high frequency of prominence systems in Austronesian languages (where the contrast is between reduced vowel and full vowel).
(iv) In Europe, the north shows more quantitative weight, while the south shows more prominence.

4 Theoretical issues

4.1 Is the onset always irrelevant for weight? The fact that onsets are irrelevant for weight seems to be a linguistic universal. There are some apparent counterexamples to this claim, to wit: some Australian languages, like Arrernte and Alyawarra (and the Amazonian language Banawá), which stress the second syllable if the first has no onset; Madimadi, which stresses the second syllable if both the first and second are light and the second has a coronal

onset; Puluwat (Oceanic; Micronesia), which stresses the first syllable starting with an /h/; and Pirahã (Mura; Brazil), which seems to have an intricate stress rule depending on presence and voicing of the syllable onset. Some examples are given in (3).

(3) Arrernte (Pama-Nyungan; Northern Territory, Australia)
ˈlelanˌtinama 'to walk along' *iˈbatja* 'milk'
Madimadi (Pama-Nyungan; New South Wales, Australia)
ˈwugaˌti 'take.IMP' *ˌwiˈridab* 'whirlwind'

The Australian languages form the largest group. For all these languages it can easily be shown that the apparent onset-sensitive stress rule is the result of a historical development (called "Initial Dropping") in which the word-initial consonant was lost, taking with it the first vowel in case it was short, while shortening the vowel in case it was long. Stress subsequently shifted to the next syllable (also in the case where a vowel still remained), creating the illusion of an onset-sensitive stress rule. Two words from Mbabaram illustrate the process.

(4) Mbabaram (Pama-Nyungan; Australia)
a. *ˈbamba* → *ˈmba* b. *ˈwaːŋal* → *aˈŋal*

If we appeal to extrametricality (cf. Chapter 14, §4.1) for the "stranded" initial vowels (these are defective syllables by Arandic standards and can be said to reject stress), then we see that these languages do not have an onset-sensitive stress rule after all, but rather show a relatively regular, quantity-insensitive, left-edged, trochaic pattern.

Madimadi is an Australian language with a truly "forbidding" surface stress pattern. Coronal second syllables seem to attract stress. A closer look, however, reveals that stress is really just located on the final syllable of the first morpheme in the word, which can be mono- or bisyllabic. In case it is bisyllabic, the medial consonant has been lenited to a coronal in a historical process. Together these phenomena create the impression of onset-sensitive stress. See Goedemans (1997) for further details.

The two remaining cases are Pirahã and Puluwat. The stress patterns in these languages are either too intricate (Pirahã) or too sketchy (Puluwat) to discuss in detail here. Suffice it to say that the patterns are either not fully described or open to more plausible reanalyses, like the ones above. See Goedemans (1996; 1998) for in-depth overviews regarding onset sensitivity in stress systems.

To date, the only case of true onset sensitivity we know of is found in Pattani Malay, where word-initial geminates attract stress, as shown in (5).

(5) Pattani Malay
ˌɟaːˈlɛː 'street/path' *ˈɟːaːˌlɛː* 'to walk' (← */bɨ+ɟalɛ/*)

This pattern, which is introduced in Hajek and Goedemans (2003), involves onset geminates, which means that here completely different structures are involved than those shown in (2).

We uphold the claim that unequivocal onset-sensitive stress rules do not exist, and follow mainstream metrical theory in focusing only on the rhyme part of the syllable when dealing with matters of weight. In that case, (2a) is always light. In some languages (2b) may be heavy while (2c) is light, in which case only vowel length plays a role. In other quantity-sensitive languages, (2c) may be heavy as well.

4.2 Quantity and vowel length. StressTyp contains at least twenty weight-sensitive languages that lack long vowels. Thus, having vowel length is not a necessary condition for having quantitative weight-effects. But what if vowel length is present? It is common to claim that in quantity-sensitive systems long vowels must be heavy **if present**. A potential counterexample to this claim is Dutch (and German), where only closed syllables act as heavy. To differentiate between the various possibilities one might adopt the idea that stress rules can "look at" different aspects of the rhyme:

(6) Options for quantity-sensitivity
– Branching of nucleus only = vowel length (2b)
– Branching of rhyme only = syllable closure (2c)
– Branching *in* rhyme = vowel length and closure (2b,c)

However, it remains to be seen whether cases like Dutch truly ignore vowel length for weight purposes. It could very well be that the alleged long vowels are not long at all, but rather tense (cf. van Oostendorp 2000).

4.3 Variability of coda weight. When codas can contribute to weight it does not mean that all consonants in coda position do so, nor that consonants that can contribute weight do so in all positions in the word.

The cross-linguistic variability in coda weight involves cases in which only certain coda consonants count as "weightful". For example, in Inga only sonorant consonants make heavy syllables, while syllables ending in an obstruent are light. The final syllable is stressed if it is heavy; otherwise stress is penultimate. Some examples are given in (7).

(7) Inga (Quechuan; Colombia)
yaˈwar 'blood' *ˈkančis* 'seven'
apaˈmuy 'to bring' *kamˈkuna* 'you.PL'

Languages with other heavy-light divisions among the set of possible codas exist as well. Quite often these involve the glottal stop. In Mam (Mayan; Guatemala), for instance, weight is assigned according to a scale (a phenomenon we find often) in which syllables with long vowels are the heaviest, followed by syllables that have a glottal stop in the coda. Syllables closed by any other consonant than glottal stop are at the bottom of the scale.

In addition, it may happen that primary stress and secondary stress differ in terms of their sensitivity to quantity (cf. Rosenthal and van der Hulst 1999). In some of these cases, consonants are weightful only in certain positions of the word. In others, the exact weight type is different for main and secondary stress.

4.4 Prominence. In type five languages, nonquantitative factors determine intrinsic syllable "weight". These factors are often referred to as prominence factors. In this section, we mention some of the different options. The first prominence factor that we discuss is (high) pitch. In languages that have contrastive high- and low-pitched syllables (i.e. tone languages), stress may be sensitive to such distinctions and, for example, be located on the leftmost or rightmost high-pitched syllable in the stress window. Look at the Northern Haida (Na-Dene; British Columbia) examples in (8), in which the last high-toned syllable is stressed, or else the last syllable.

(8) Northern Haida
ˌguudingˈee 'giant purple urchin'
ˌʔadlaˌdajanˈdáálˌgang 'jump.up.ITER.along.PRES'

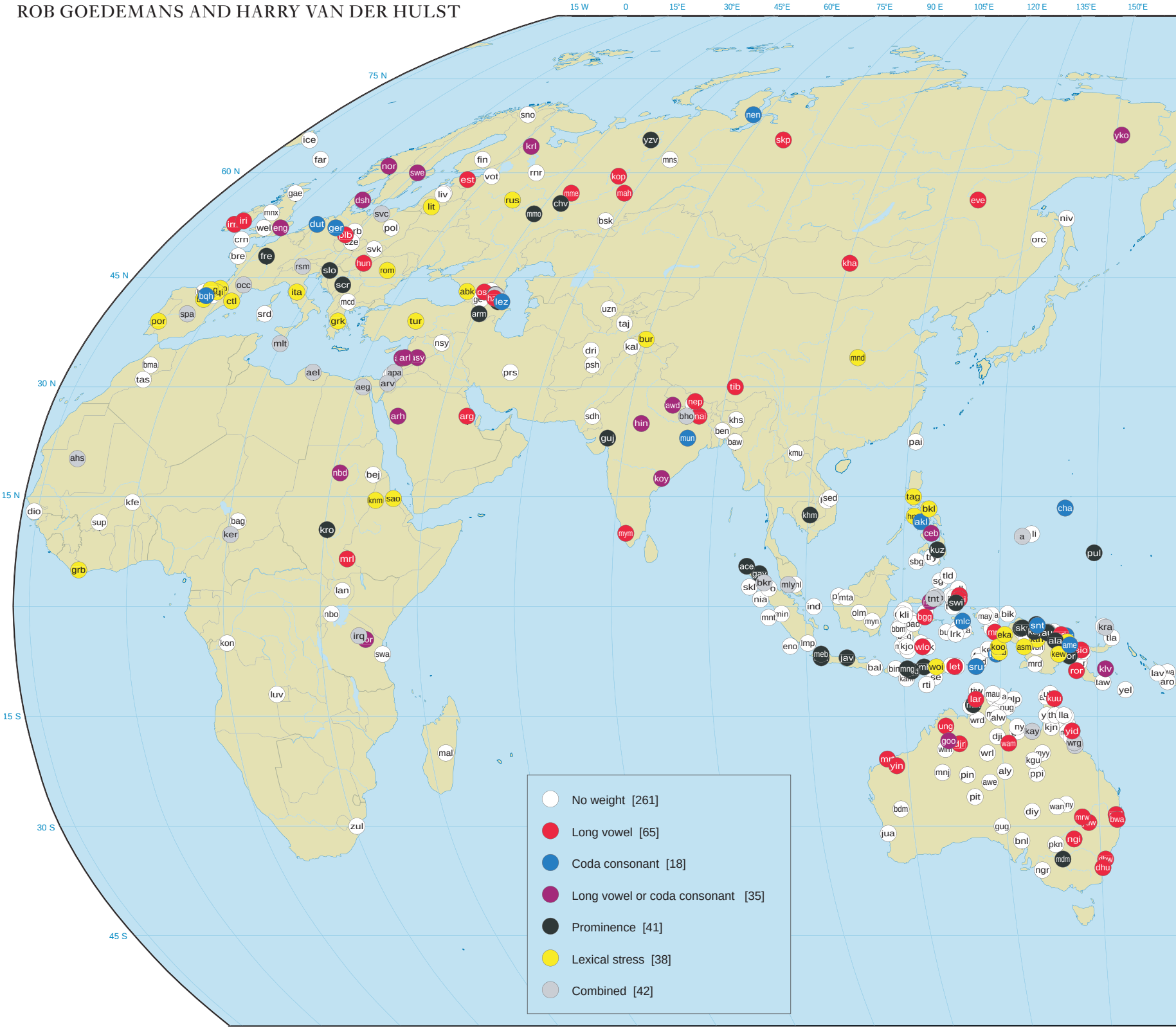
Another prominence factor concerns vowel aperture, or more generally vowel quality. If aperture is relevant, more open vowels will count as heavy, as opposed to closed vowels. The Yindjibarndi (Pama-Nyungan; Western Australia) examples in (9) show initial stress unless the second syllable contains a low long vowel.

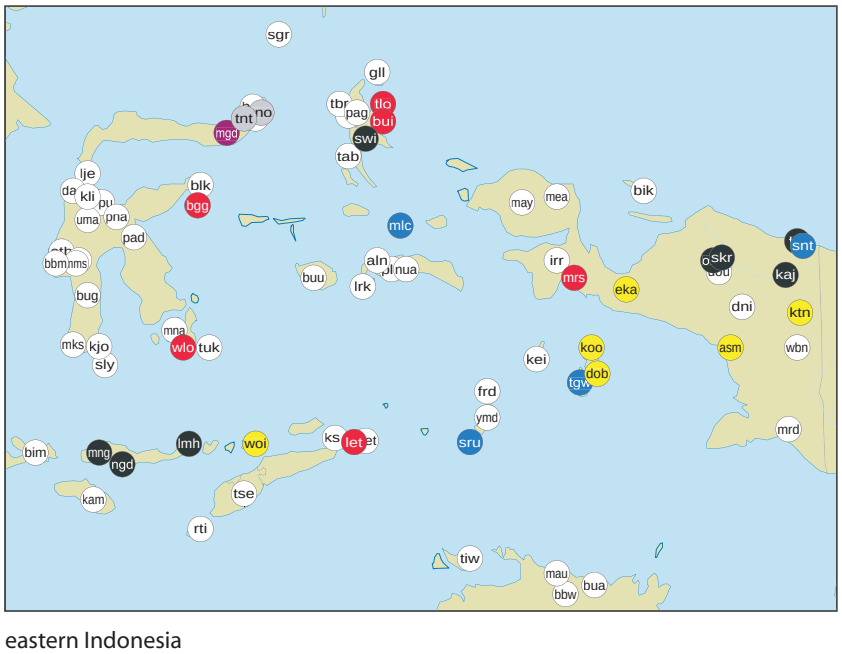
(9) Yindjibarndi
ˈmartuurˌraa 'twilight' *nyiˈlaarti* 'native mead'

This behaviour reflects a general tendency among prominence factors. Many of these divide up syllables such that the more sonorous ones are heavy while the others are light (recall Inga, which is basically also a prominence system in which only sonorant codas make syllables heavy). If overall vowel quality is relevant, the opposition typically is reduced (light) as opposed to full (heavy) vowels. For examples, see the Chuvash words in §2.5. More on prominence distinctions among different codas and vowels can be found in Zec (1988) and Kenstowicz (1994).

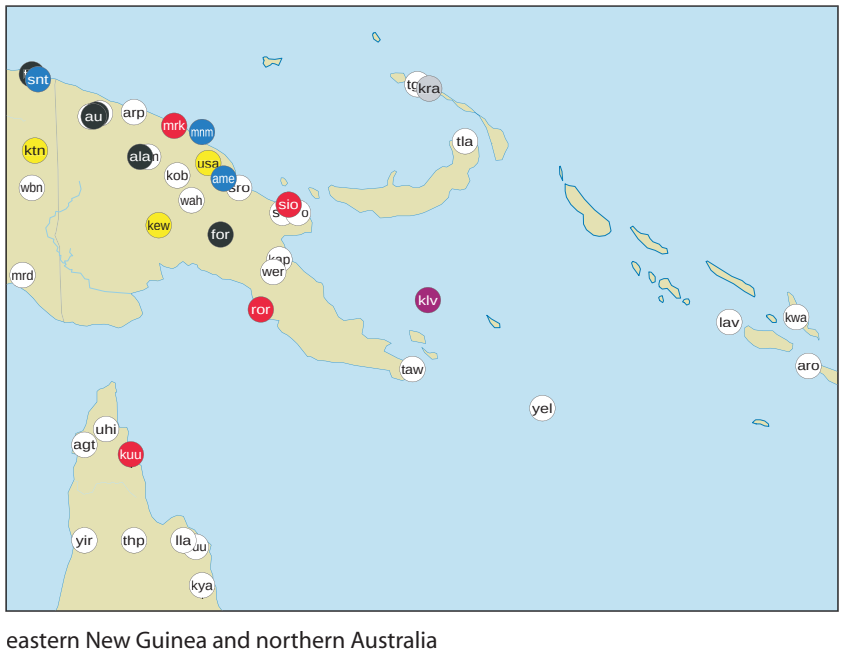
16 Weight Factors in Weight-Sensitive Stress Systems

ROB GOEDEMANS AND HARRY VAN DER HULST

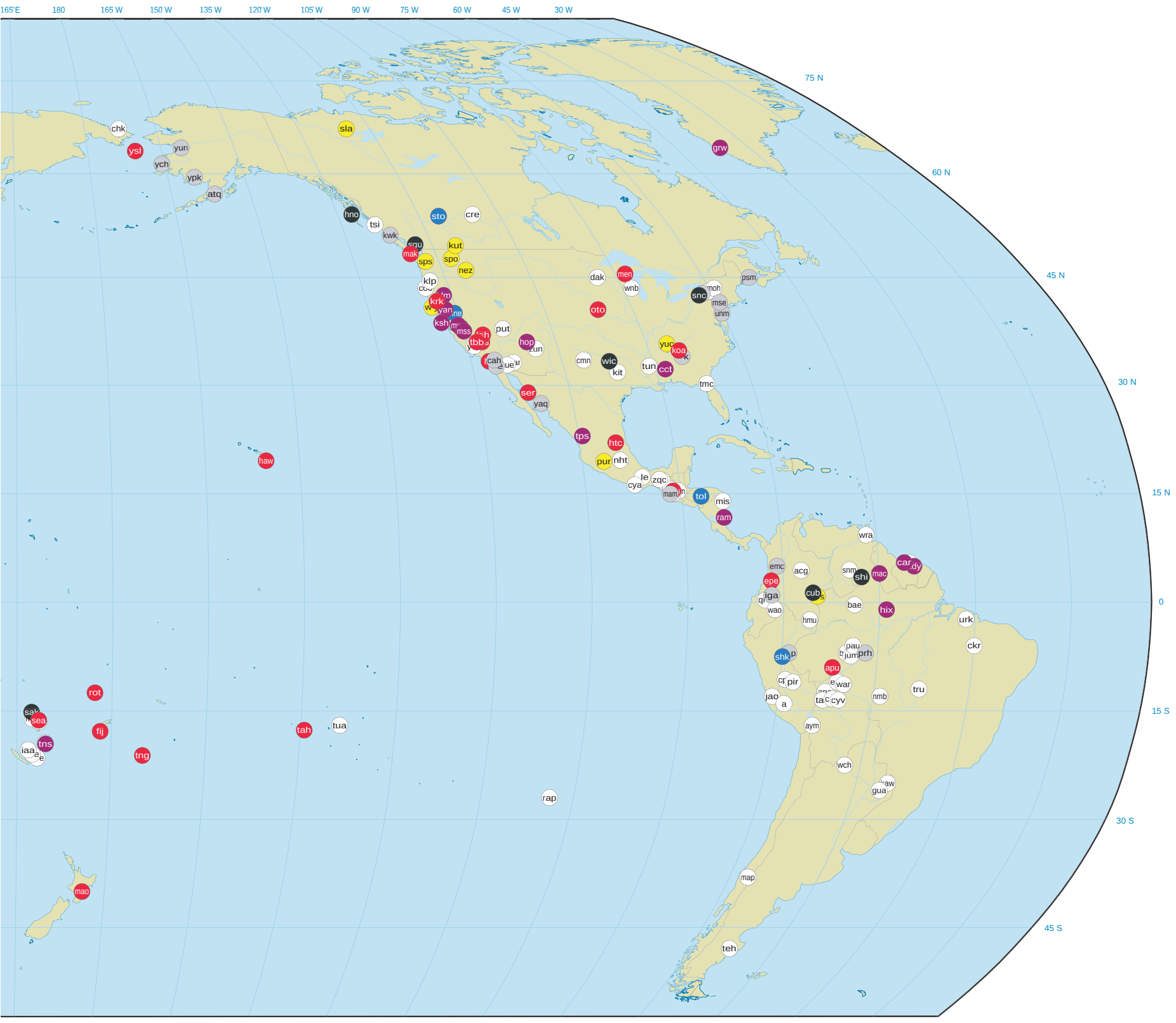


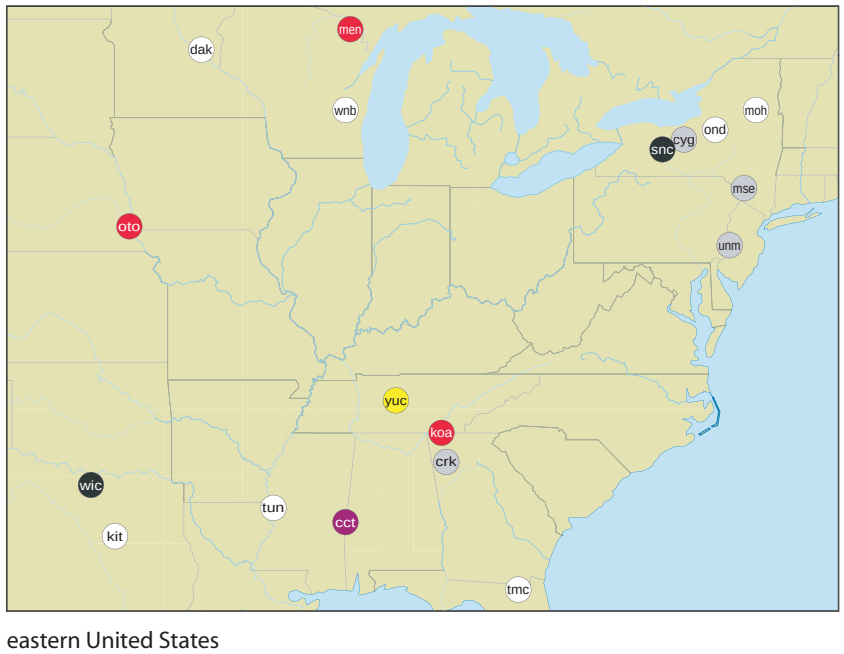


eastern Indonesia



eastern New Guinea and northern Australia





eastern United States

Mesoamerica

17 Rhythm Types

ROB GOEDEMANS AND HARRY VAN DER HULST

1 The separation of primary and secondary stress

Assuming that the phonemes that make up words are organized into syllables, modern linguistic theory claims that **rhythm** is a manifestation of the fact that syllables are further grouped into constituents called **feet**, which are usually binary groupings of syllables. Standard Metrical Theory takes the grouping of syllables into feet as fundamental for the assignment of bounded primary stress. StressTyp, however, is based on a separation of the treatment of **primary** and **secondary stress**, and the latter (but not the former) primarily has to do with rhythm-based stresses. Such a separation has been proposed and defended on theoretical grounds by van der Hulst (1996). We offer several arguments here.

First, in our discussion of bounded systems (Chapter 15, §2.1) we have seen that weight-sensitive primary stress locations within the two-syllable window can take four different forms (at each edge of the word). **Rhythm types**, however, fall into just two categories, due to the fact that a rhythmic foot cannot contain two heavy syllables (because every heavy syllable constitutes a rhythmic beat). This difference, then, may make it necessary to describe primary and secondary stress patterns separately.

A second argument for the separation can be found in systems in which secondary stress starts at the edge that is opposite to the location of primary stress; we refer to such systems as *polar*. An example is given in (1).

(1) Garrwa (Garawan; Northern Territory, Australia): initial main stress, secondary stress on the penultimate and every second syllable before it.
('nari)ɣin(ˌmukun)(ˌjina)(ˌmira) ‘at your own many’

In Table 1 we see exactly how many languages this argument is based on (data in Tables 1–4, and Figure 1, are based on the *complete* StressTyp sample of 510 languages).

Table 1 Frequencies of languages with left- or right-edged main stress, cross-spliced with direction of footing

	Starting edge of footing (secondary stress assignment)	
Main stress domain	**Left**	**Right**
Left	63	12
Right	27	53

In total, 39 languages (8 per cent of all the languages in StressTyp) assign main and secondary stress from opposite directions. It would seem that the subset of languages for which we need separate direction parameters for main stress and rhythm is large enough to provide serious support for the separation of primary and secondary stress.

A third argument for the separation of primary and secondary stress is formed by languages in which the rules for main stress and rhythm do not agree in their usage of weight, or, more intricately, in the way in which they use it. In Kayardild, for instance, main stress is initial, regardless of weight, but long vowels receive a secondary stress, whereas short vowels do not (simplified for expositional purposes).

(2) Kayardild (Tangkic; Queensland, Australia)
'kunja ‘little’ *'maˌlaa* ‘sea’

Table 2 shows the number of languages in which main stress is weight-sensitive and secondary stress is weight-insensitive, or vice versa. The last column in Table 2 shows the languages in which main and secondary stress disagree in their definition of what constitutes a heavy syllable. In such languages, closed syllables may, for instance, be heavy for stress but not for rhythm.

Table 2 Mismatches in quantity-sensitivity between primary and secondary stress (weight-sensitive languages for which we have no information on rhythm were not taken into account)

Both weight-sensitive	40	41%
Primary weight-sensitive, Secondary weight-insensitive	32	33%
Primary weight-insensitive, Secondary weight-sensitive	17	18%
Both weight-sensitive, but differently	8	8%

In total, we need separate quantity statements for primary and secondary stress in 57 languages (11 per cent of all the languages in StressTyp). All in all, 17 per cent of the languages in StressTyp support the separation hypothesis (Tables 1 and 2 combined, ten overlapping languages).

A fourth argument for separating primary and secondary stress assignment lies in the fact that whereas lexical marking is quite normal for primary stress, even in systems that have dominant rule-governed locations, secondary stresses are never a matter of lexical marking. In this statement, we ignore so-called “cyclic stresses”, i.e. secondary stresses that correspond to primary stress locations in embedded morphemes in complex words.

2 Defining the values

For the languages for which we have information on rhythm, the following types are shown on the map.

1. Trochaic: left-hand syllable in the foot is strong	153
2. Iambic: right-hand syllable in the foot is strong	31
3. Dual: system has both trochaic and iambic feet	4
4. Undetermined: no clear foot type	37
5. Absent: no rhythmic stress	98
total	323

2.1 Trochaic rhythm. The prototypical rhythm pattern stresses every odd syllable from the left in languages with initial main stress, and every even syllable from the right in languages with penultimate main stress. This is achieved with trochaic feet, as in Ono (Trans-New Guinea; Papua New Guinea): *('mesi)(ˌkene)* ‘you will sit’, *('ari)(ˌmage)(ˌake)* ‘he always goes’.

2.2 Iambic rhythm. Languages with iambic rhythm typically stress even syllables from the left or odd syllables from the right, as in Mapudungun (Araucanian; Chile): *(e'lu)(muˌyu)* ‘give us’, *(ki'mu)(faˌlu)(wuˌlay)* ‘he pretended not to know’.

2.3 Dual rhythm. This type is difficult to find. One needs to look at all the data of a language carefully to discover that one foot type is not sufficient to cover all cases. An example of such a language is Yuwaalaraay (Pama-Nyungan; New South Wales, Australia): *(gaˌlu)(ma'ya)* ‘grandson’, *('gidul)(ˌgara)* ‘snail’.

2.4 Undetermined rhythm. These languages do have secondary stresses, but the patterning falls into none of the above types. Languages like Malayalam (Dravidian; India) which place secondary stress on long vowels only fall under this type: *'pukavaɳʧi* ‘train’, *'mutaˌlaaʃi* ‘boss’.

2.5 No rhythm. Languages that have main stress but no rhythm abound. An example is Gorowa (Southern Cushitic; Tanzania): *giram'booda* ‘snuff’, *oro'mila* ‘because’. It remains an open question whether the “no rhythm” languages do have a foot-type grouping but lack a clear phonetic manifestation of this grouping; in some cases such “silent” grouping can be detected because it conditions other aspects of the phonology of a language.

3 Geographical distribution

We observe the following tendencies:

(i) Iambic rhythm occurs mostly in North and South America.
(ii) South America and Australia always seem to have clear rhythmic patterns.
(iii) Africa, on the other hand, shows little evidence for rhythmic patterns.

As was noted in Chapter 14, we find an unequal division of languages over the two foot types that we use to describe the rhythmic properties of stress systems. In general, languages prefer the trochaic foot over the iambic foot. Note again that we divide languages into trochaic or iambic systems on the basis of their *secondary stress* properties. Unfortunately, that means that many languages in StressTyp are left out of the equation, simply because we do not have any secondary stress data for these languages. Statistics for the languages for which we do know something about secondary stress can be found in Figure 1.

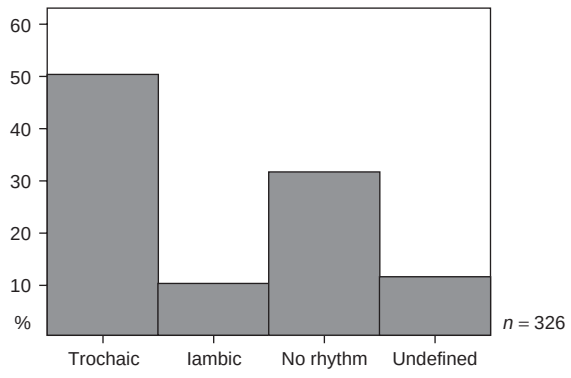


Figure 1 Rhythm types

The hugely skewed distribution between trochaic and iambic systems in Figure 1 is hard to miss even in a cursory look at the typology of stress systems. Though it has never been quantified in this way before, the scarcity of iambic systems has already led some researchers on the path to total denial of the iamb's existence as an abstract metrical entity. Hayes (1995) dispenses with all quantity-insensitive iambs, while among the most drastic of these studies, van de Vijver (1998) claims that *all* iambic systems can be reanalysed as trochaic ones (cf. van der Hulst 1999b for discussion). In StressTyp we choose no sides in this obviously highly theoretical debate. We will continue to offer the iambic foot as an option, because there are enough languages, in our view, whose stress patterns are clearly more straightforwardly described with the help of these iambic feet.

4 Theoretical issues

4.1 Correlations between rhythm type and edge. A correlation that has been noted many times in the past concerns the foot type used in secondary stress assignment and the edge at which this assignment starts. The claim is that languages with an iambic rhythmical pattern assign rhythm from left to right, as in Mapudungun in §2.2 (cf. Kager 1993). An example that makes this clear is given in (3).

(3) Mapudungun
(e'lu)(a,e)new 'he will give me'

Rhythm assignment from right to left would have given us *elu'ae,new*. (In this example we note that a left-over syllable, here on the right side, is not footed and hence has no rhythmic beat; this is a quite common phenomenon, and it has been argued that "degenerate", i.e. monosyllabic, feet are uncommon or universally banned in quantity-insensitive languages.)

The reverse claim, that languages with trochaic rhythm start footing at the right edge, is not commonly deduced from the observation with respect to iambic languages. Linguists seem hesitant to claim anything in this case. We present the data for both types in Table 3.

Table 3 Differences in preference for direction of footing for both foot types

	Starting edge of footing	
Foot Type	Left	Right
Trochaic	83	58
Iambic	22	8

We observe that the correlation between iambic rhythm and left-to-right footing is indeed quite high. It turns out to be even higher when we remove some "suspicious" cases from the list of right-to-left iambic languages. It appears that in many of these languages the choice between iambic and trochaic feet is not easily made, and where the descriptive source mentions that the language is iambic, this choice is often based on superficial evidence. For only three of these cases can we be quite certain that rhythm is iambic.

For trochaic patterns, we note that most trochaic languages start foot assignment at the left instead of the right edge. Yet the difference is not large enough to speak of a clear correlation in this case. We may conclude, therefore, that the choice to state a correlation between foot type and direction for iambic systems is justified only in hindsight, by the StressTyp data.

4.2 Correlation between rhythm type and relevance of weight. Hayes (1995) suggests as another correlation, this time between rhythm type and weight-sensitivity, the so-called **iambic–trochaic law**. According to this "law" weight-insensitive languages (grouping units of equal weight, so to speak) are (preferred to be) trochaic. The reverse claim, that weight-sensitivity leads to iambic grouping, is less obvious, although one idea is that weight-differences based on quantity favour iambic patterns, while weight differences based on "intensity" favour trochaic patterns. However, it is not clear what intensity is (prominence?); moreover, many weight-sensitive trochaic patterns are based on quantity (cf. Rice 1992, Kager 1993, and van der Hulst 1999b for discussion). The following table indicates the relationship between the presence of weight-sensitivity and the rhythmic foot type:

Table 4 Correlation between weight and foot type in StressTyp

	Weight	
Foot Type	no	yes
Trochaic	103	51
Iambic	11	17

4.3 Iterative and noniterative secondary stress. So far we have assumed that the rhythmic structure of a word comprises the whole word. However, rhythm assignment can be 'noniterative', which means *only one* secondary-stress foot is formed, as in the Ngalkbun (Australian) example in (4); here only one secondary stress occurs at the opposite edge from the main stress, though the word is long enough to allow more.

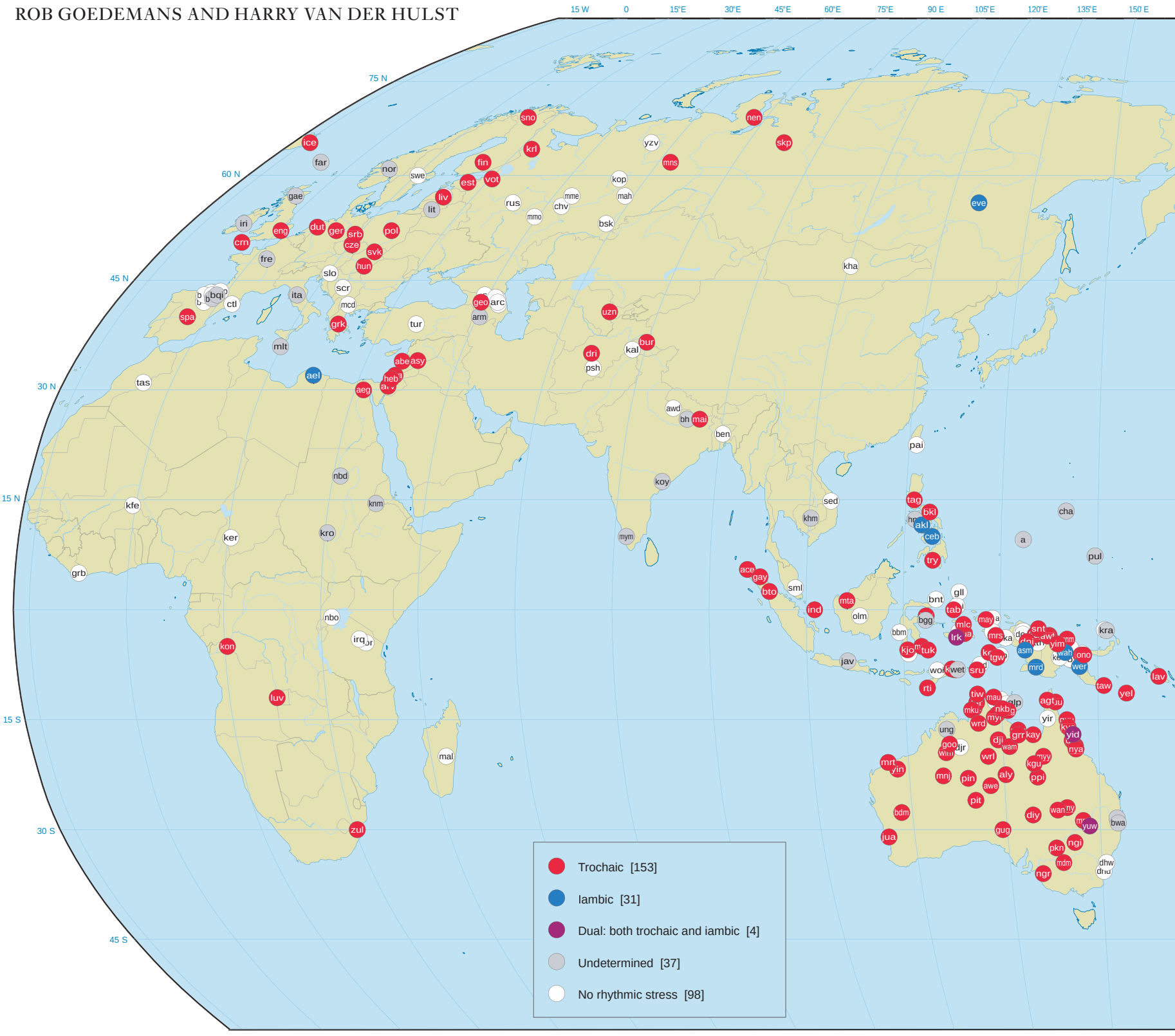
(4) Ngalkbun
ˌŋaʔyenjyenjtjuˈŋiyan 'I will talk'

StressTyp reveals that the iterative option is clearly favoured. 70 per cent of the languages for which we have secondary stress information choose this option, while only 18 per cent reveal noniterative rhythm (no iterativity information for 12 per cent). For the sake of completeness we mention here that six of the iterative secondary stress languages show **ternary stress** (the rhythm beat is followed, or preceded, by *two* unstressed syllables); cf. Rice 1992.

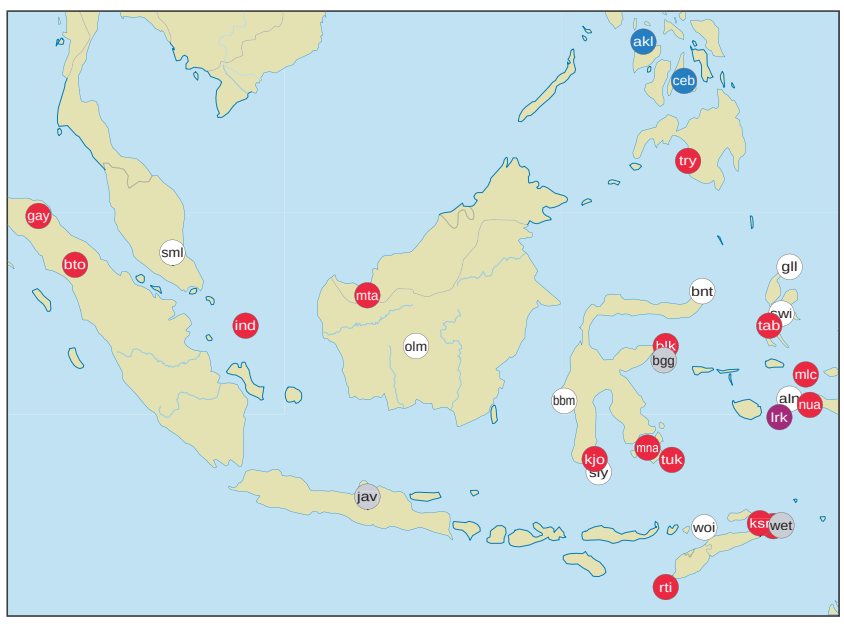
4.4 Foot typology. The precise grouping of syllables into feet is subject to theoretical discussion. In the classical theory, one heavy and one light syllable could be grouped together in a foot, but in later versions of the theory (the so-called Revised Theory) heavy syllables form feet by themselves. A choice on this issue does not, however, influence whether a language is classified as iambic or trochaic. We refer to Hayes (1995) and van der Hulst (1999b) for discussion of this issue.

17 Rhythm Types

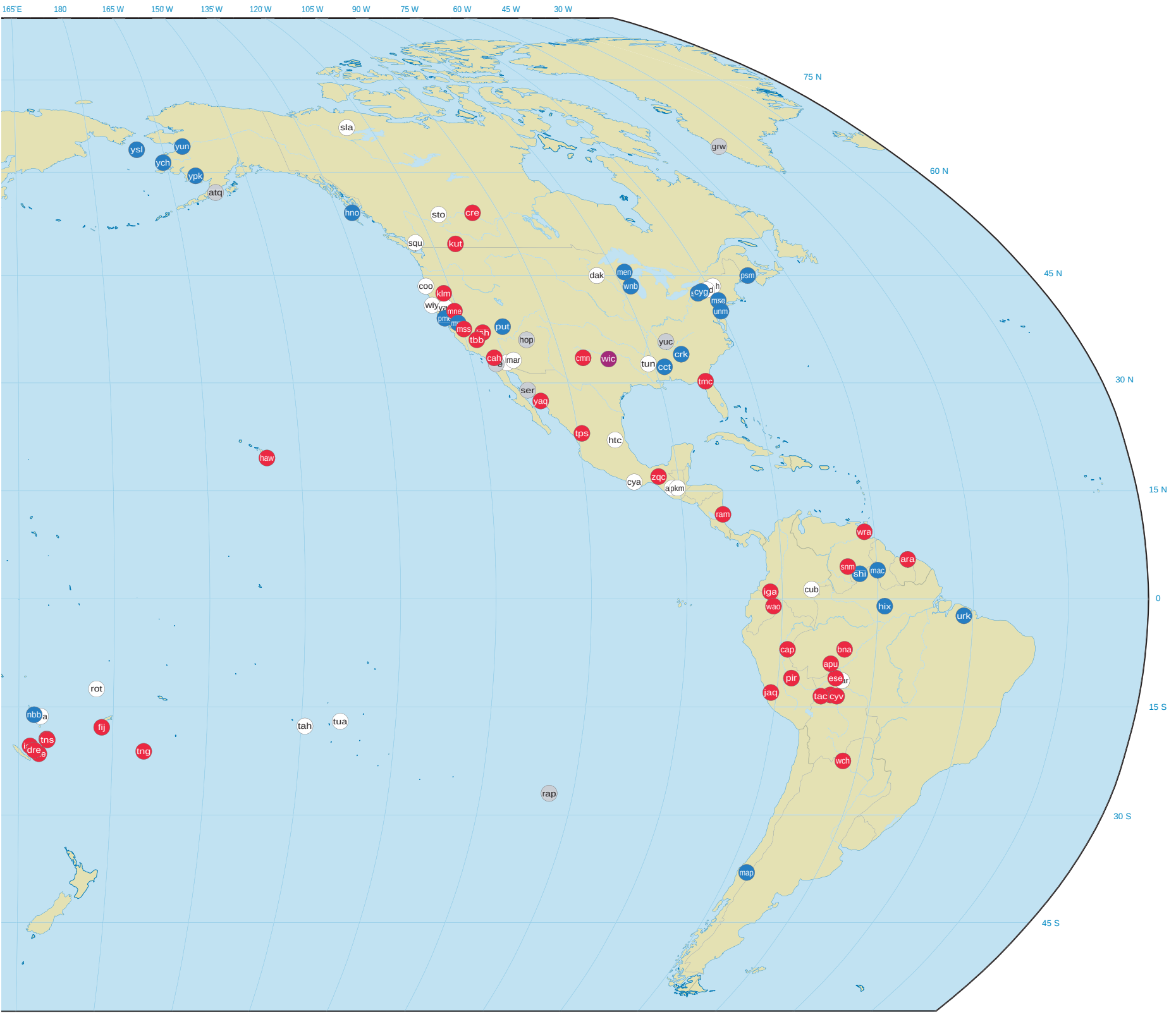
ROB GOEDEMANS AND HARRY VAN DER HULST

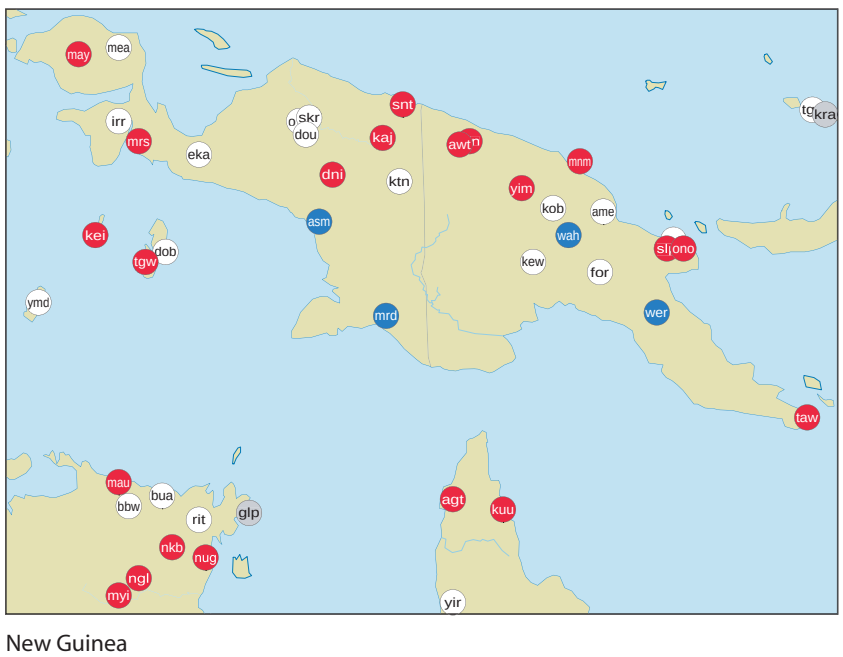


Europe

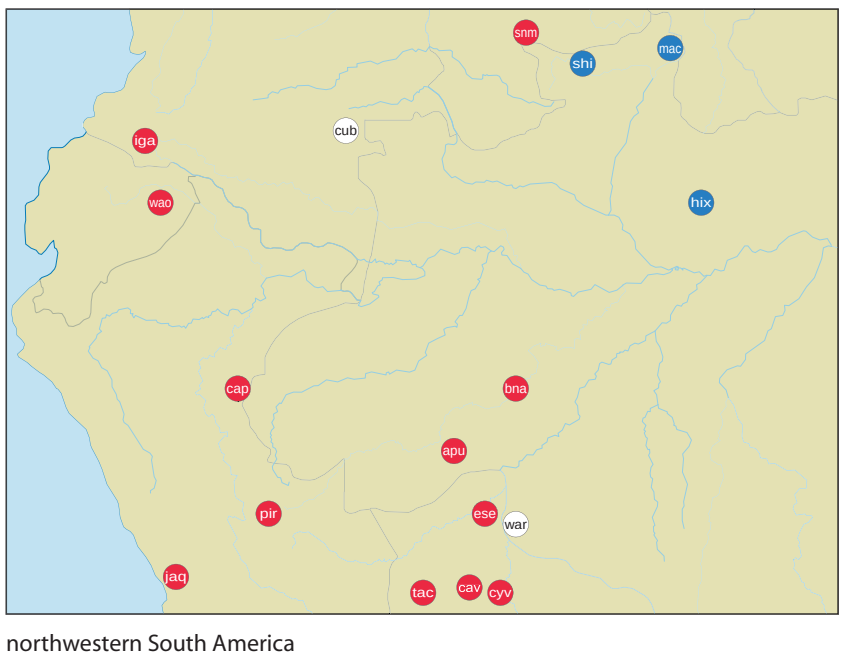


Sumatra, Borneo, and surrounding areas





New Guinea



northwestern South America

18 Absence of Common Consonants

IAN MADDIESON

1 Introduction

In the chapter on the size of consonant inventories (Chapter 1) the kinds of sounds that are most often found in the consonant inventories of the world's languages were briefly discussed. In this chapter the focus is on the absence of some of the classes of sounds which typically occur in the world's languages, and on where these absences occur. Absences of three classes of consonant sounds—bilabials, fricatives, and nasals—which occur in the vast majority of languages will be discussed. Each of these terms will be defined before discussion of the geographical distribution of the pattern. Absence of a fourth class of common consonants, laterals, is discussed in Chapter 8.

1.	All present	502
2.	No bilabials	3
3.	No fricatives	48
4.	No nasals	10
5.	No bilabials or nasals	2
6.	No fricatives or nasals	1
	total	566

Of the languages whose consonant inventories were analysed for this chapter, 502 (or 88.7 per cent) have at least one representative of each of the three classes of consonants being discussed, and only 64 (or 11.3 per cent) show the absence of one or more of them. The languages with all the classes present are classed into an "all present" category. A striking fact is that not a single language on the Eurasian landmass shows an absence of any of these consonant classes. The absences are found predominantly in the Americas, Australia, Melanesia, and the other Pacific islands.

2 Absence of bilabials

Bilabials are the class of sounds made by bringing the two lips towards each other, especially as in the sounds commonly represented by the letters *p*, *b*, *m* (for example in the English words *pub* and *mob*). There are also bilabial fricatives, written in phonetic transcription with the symbols /ɸ, β/, but these are less common than the stops /p, b/ and the nasal /m/. The vowel-like consonant represented by /w/ (e.g. in the English word *wood*) is not considered bilabial, as the movement of the lips is forward, as in a rounded vowel, rather than primarily being a movement of the two lips towards each other in the vertical plane. One or more consonants at the bilabial place of articulation occurs in the consonant inventories of virtually all languages, but in a few languages this place is absent. In contrast, all known languages have at least one consonant in which the tip or blade of the tongue makes a closure against or behind the upper teeth (dental or alveolar), and only one language, Dumo (Sko; Sandaun, Papua New Guinea), is currently known to have no velar consonants. Among these three major place categories—bilabial, dental/alveolar, and velar—the bilabial is the one which is most frequently missing altogether. Even then, its absence is far from common.

Just five of the languages in the survey lack a bilabial consonant. All of the languages concerned are indigenous to North America, but for the most part are neither geographically nor genealogically particularly close to each other. In four of these languages the absence of bilabial consonants is straightforward. This is the case for Tlingit and Chipewyan, two languages in different branches of the Na-Dene family, the Iroquoian language Oneida, and Wichita, a member of the small Caddoan language family. In the fifth language the absence of bilabials is partly a question of interpretation. Eyak (Na-Dene; Alaska), a distant relative of Tlingit and Chipewyan, is considered here as a language without bilabials even though the bilabial nasal sound [m] is heard in this language. This sound only occurs before a nasalized vowel, and the sound [w] is not pronounced before a nasalized vowel; therefore the bilabial nasal is considered the variant of /w/ in the position before a nasalized vowel and not a member of the inventory of distinct consonants of Eyak. The way that the sounds [w] and [m] are distributed in this language is part of the reason it is also considered not to have nasal consonants, as will be discussed below. Oneida is also classed as lacking nasals, and so forms with Eyak a special subcategory of the "no bilabial" languages. Another language of the American far north, Aleut, has no bilabial stop sounds or fricatives but does have both voiced and voiceless bilabial nasals. This pattern is rare, as nasal consonants usually occur in a language's consonant inventory only if there are non-nasal consonants at the same place of articulation.

3 Absence of fricatives

An absence of fricatives is far more frequent than an absence of bilabials. Fricatives are the sounds that are created when air from the lungs passes through a sufficiently narrow constriction on its way out of the mouth to set the air into a noisy turbulent flow pattern. The sounds at the beginning of the English words *fan*, *van*, *thin*, *that*, *sign*, *zone* are fricatives. The sound represented by *h*, as at the beginning of *hope*, is not considered a fricative for the purposes of this chapter. Although this sound is quite often described as a glottal fricative it is not produced with a constriction similar to other fricatives, and it frequently has a different pattern of distribution in words from the true fricatives. For example, in English (and other Germanic languages) /h/ can only appear at the beginning of a syllable directly before a vowel and not at the end of a syllable nor in a sequence of more than one consonant. True fricatives, such as /s/ and /f/, occur both at the beginning and at the end of syllables and can co-occur in these positions with other consonants, as in words such as *flan*, *sleep*, *wolf*, *pulse*.

An absence of fricatives is far more common than an absence of bilabials, being noted in 49 (or about 8.7 per cent) of the languages in the sample. The great majority of these languages are in Australia, with other notable clusters in New Guinea and in the interior of South America. Outside of these areas there are only a few sporadic examples of languages without fricatives, such as Kiribati and Hawaiian (both Austronesian), the Nilo-Saharan languages Dinka and Lango, spoken in the Sudan and Uganda respectively, and the one surviving Great Andamanese language (also known as Pucikwar), as well as Aleut, as mentioned above.

The absence of fricatives is broadly characteristic of the languages of Australia. Only four of the Australian languages in the sample have any fricatives, against twenty-nine which have none. Furthermore only one of the four languages with fricatives, Kala Lagaw Ya (Pama-Nyungan; Torres Straits Islands), has fricatives at the most common place of articulation and of the most common type, namely the sibilant sounds represented by /s, z/. Most languages with only one fricative have a voiceless sibilant of the type usually represented by the letter *s*, but the three other Australian languages with fricatives have just one voiced fricative each. This is unusual, first because fricatives are more commonly voiceless (see Chapter 4), and secondly because these fricatives are of less common types, namely velar in Tiwi (Northern Territory) and Maung (Yiwaidjan; Northern Territory), and an alveolar nonsibilant in Ngiyambaa (Pama-Nyungan; New South Wales) reminiscent of the sound represented by *th* in the English word *these*. The rarity of fricatives in Australian languages is one of several properties of their phonological inventories, such as their characteristically smaller

than average vowel inventories and their lack of voicing contrast in stops (as discussed in Chapters 2 and 4), which make these languages stand out from others. The relatively frequent absence of fricatives from Papuan languages as well (for example in Yimas, Gadsup, Wahgi, and Yelî Dnye) establishes a degree of typological similarity between language groupings which are not known to be historically related and have certainly not been in recent contact, but between which an ancient connection is not implausible.

4 Absence of nasals

Nasals are consonants in which there is a closure in the mouth cavity and the entire flow of air is directed through the nose. The sounds associated with the letters *m*, *n* in English words like *man* as well as the sound at the end of words like *sing* are nasals. A total of thirteen languages in the sample are listed as having no nasals in their consonant inventories. Some of these languages, such as Quileute (Chimakuan; Washington State), Rotokas (West Bougainville; Papua New Guinea), and Pirahã (Mura; Brazil), make no systematic use of nasality in their sound system at all; the last two have especially small phoneme inventories overall. The majority of these languages, however, do make use of nasality, but it patterns in such a way that simple nasal consonants do not need to be considered contrastive segments. The case of Eyak has been mentioned above. Although both [m] and [n] sounds are heard in this language, [m] is a variant of /w/ before a nasalized vowel and [n] is the variant of /l/ in the same position. Oneida is also analysed here as a language with no nasals in its consonant inventory, although the sound [n] occurs in this language and other interpretations might treat it as a basic consonant. As already noted, Eyak and Oneida are the two languages classed as having neither bilabials nor nasals. Maxakali (Macro-Ge; Brazil) has no fricatives as well as no nasals and therefore forms a unique class in the sample.

The majority of the languages considered to have no nasal consonants make a distinction between oral and nasalized vowels after consonants such as voiceless plosives. But voiced plosives only occur before oral vowels and simple nasal consonants occur before nasalized vowels. In this situation the nasals can be considered variants of the voiced plosives. This pattern is found, for example, in Klao (Kru, Niger-Congo; Liberia) and Andoke (isolate; Colombia). In a number of these languages a type of plosive which consists of a nasal part and an oral part, similar to what is heard in the middle of the English words *timber*, *thunder*, *anger*, is also involved in the pattern. Sounds of this sort considered as single segments are called **voiced prenasalized stops**. In Maxakali there are both prenasalized stops and nasalized vowels in the inventory. Before a nasalized vowel simple nasal consonants occur as the variants of the prenasalized stops in this position. In Cubeo (Tucanoan; Colombia) voiced plosives, voiced prenasalized stops, and simple nasals are all heard, but each type of sound can only occur in particular positions. The voiced plosive sounds occur between two oral vowels, the voiced prenasalized stops occur between a nasalized and an oral vowel, and plain nasal consonants occur before a nasalized vowel. These types of sounds therefore do not contrast with each other and only one of them—in this case the voiced plosive type is chosen—is taken as belonging to the set of distinctive consonants of Cubeo.

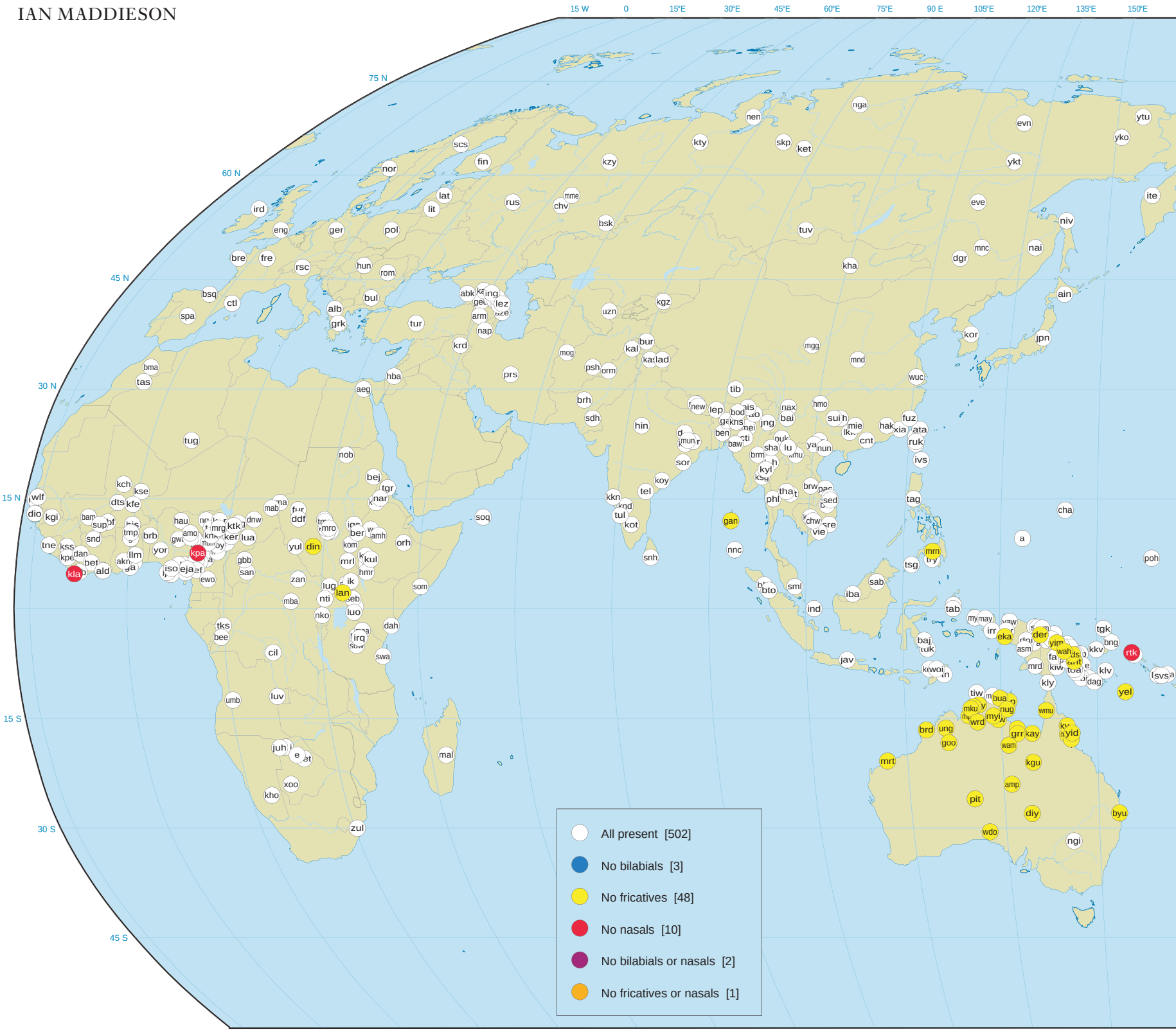
As these examples make clear, although there are a number of languages which can be analysed as having no nasal consonants, extremely few of the world's languages fail to make use of nasality either as a part of their consonant system or as part of their vowel system. The ability to direct the flow of air through the nose is used in one way or another in virtually all human languages.

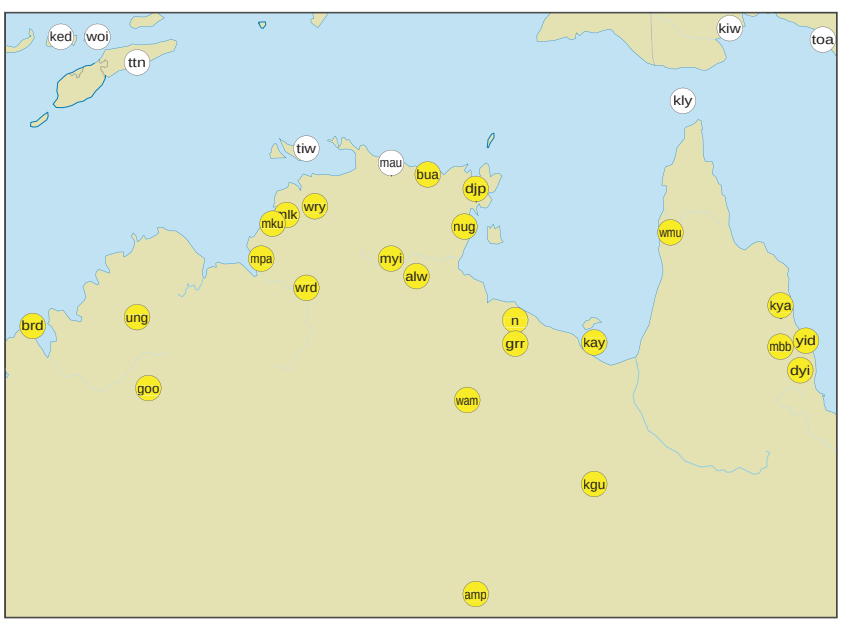
5 Summary

The three types of absences of familiar consonants analysed in this chapter show interesting differences. All three show areal concentration, but it is far more intense for the absence of fricatives than for the other two cases. The absence of fricatives and of bilabials is often straightforward, whereas the absence of nasals usually depends on analytical choices made by the linguist. In particular the property of frication per se is itself generally absent from the languages with no fricatives, whereas the property of nasality is usually present even in those languages analysed as having no nasals.

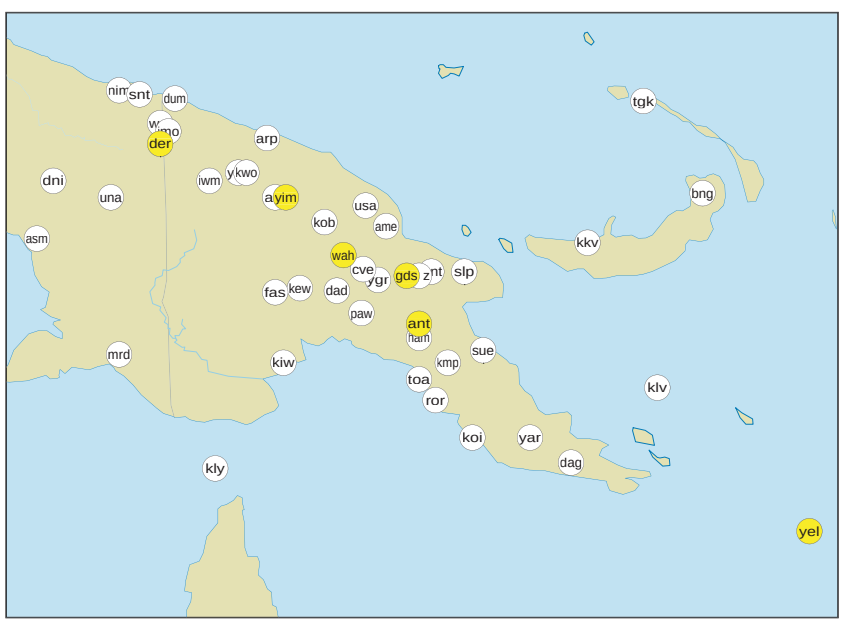
18 Absence of Common Consonants

IAN MADDIESON

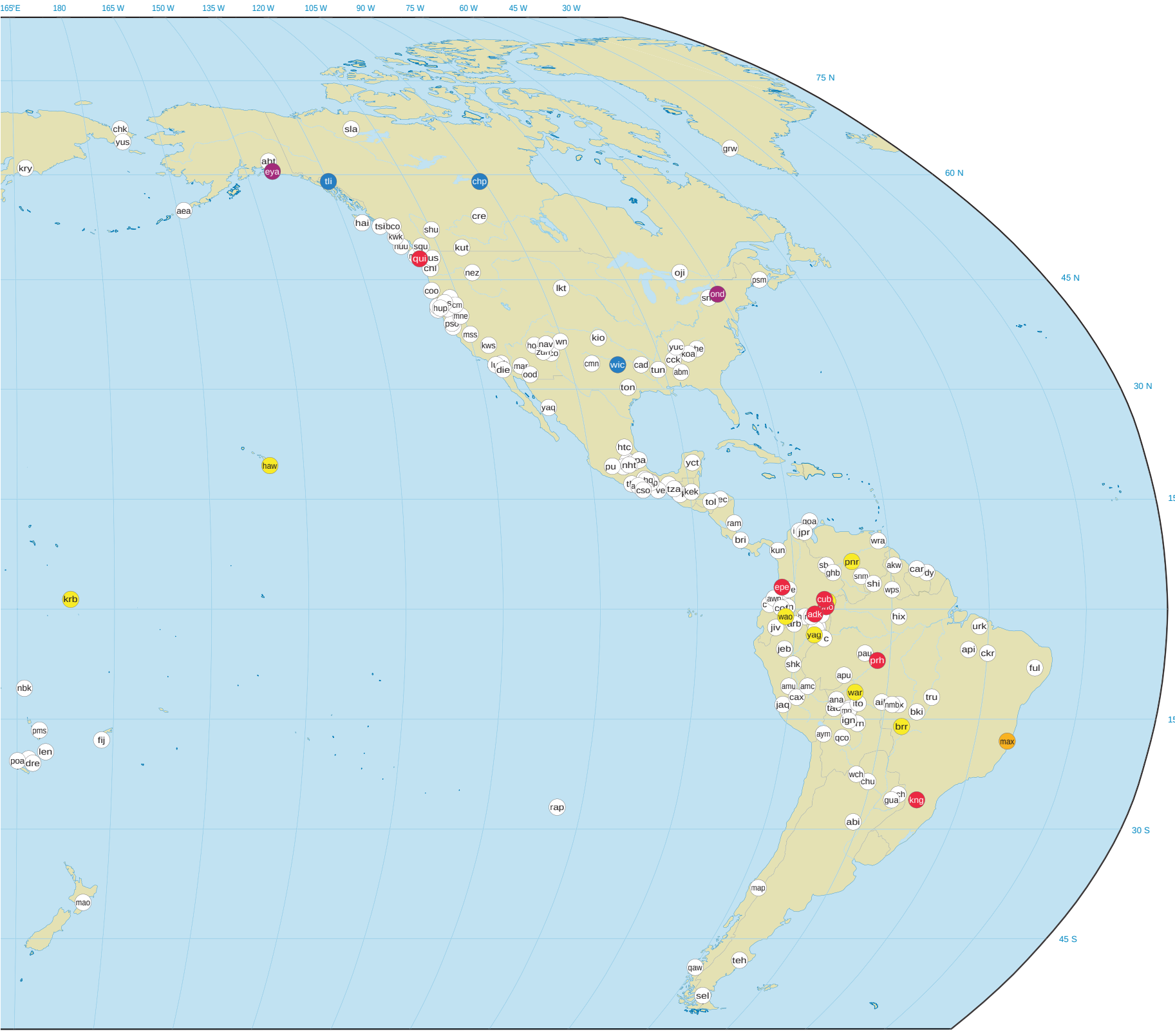




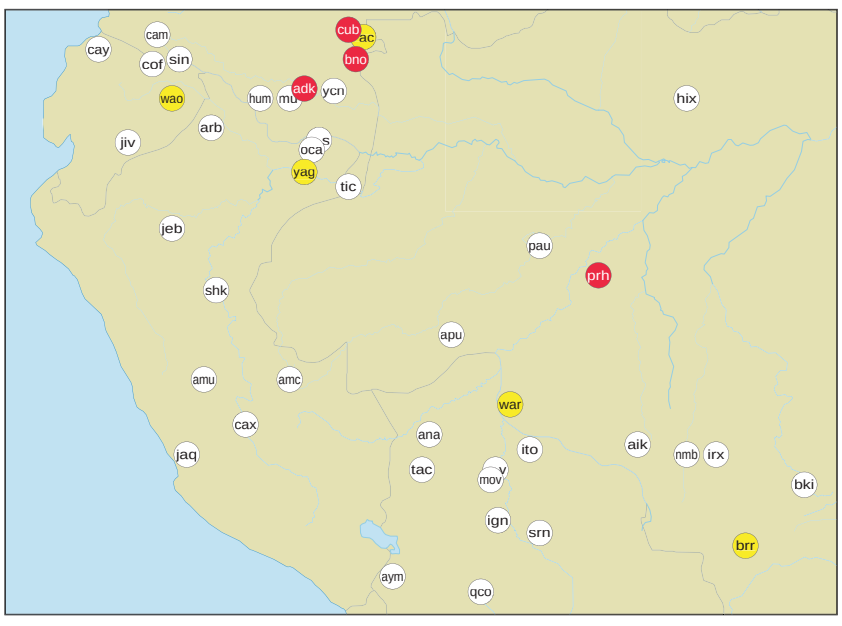
northern Australia



eastern New Guinea



North-West Coast (Alaska and British Columbia)



northwestern South America

19 Presence of Uncommon Consonants

IAN MADDIESON

1 Introduction

In this chapter the worldwide pattern of occurrence of four rather unusual classes of consonants will be discussed. These four classes are clicks, labial-velar plosives, pharyngeals, and dental or alveolar nonsibilant fricatives. This last class includes sounds similar to those represented by the *th* in English words such as *this* or *thing*; for brevity they will often be referred to as "th" sounds in the text. Each of these classes of consonants occurs in relatively few languages, but they exemplify different patterns of geographical distribution, some being concentrated in particular areas while others are dispersed.

The occurrence of these consonants in the consonant inventory as discussed in Chapter 1 was examined in a sample of 566 languages. Since each of the four types of consonants being discussed is relatively rare, the great majority (448, or 79.2 per cent) of the languages do not have any of them. However, viewing the world on a large scale, at least one of these special kinds of consonants tends to be found in at least some language in all major areas.

1.	None	448
2.	Clicks	9
3.	Labial-velars	45
4.	Pharyngeals	21
5.	"Th" sounds	40
6.	Clicks, pharyngeals, and "th"	1
7.	Pharyngeals and "th"	2
	total	566

2 Clicks

The rarest of the four classes are the clicks, which occur in just ten (or 1.8 per cent) of the languages in our sample. Click sounds are used in many cultures for nonlinguistic communication, and a separate chapter (Chapter 142) is devoted to these uses, but they occur as regular linguistic sounds incorporated into ordinary words only in some languages of southern and eastern Africa. Clicks are sometimes regarded as particularly exotic speech sounds, but their production is not especially unusual. The mechanism for making clicks is illustrated in Figure 1. The key is the making of closures at two different locations in the mouth, so that a small volume of air is enclosed in the space between the two. One closure is always made by the back of the tongue against the back part of the palate. The other closure is made by the lips or by the tip or blade of the tongue, as is shown in the first panel of the figure. The enclosed air space is marked by shading. The centre part of the tongue is then lowered, enlarging the air space and creating a partial vacuum between the two closures, as illustrated in the second panel. The closure at the front of the mouth is then released as indicated in the third panel, and because the pressure of the air inside the mouth is lower than the pressure outside, the air abruptly rushes into the mouth, creating the click noise.

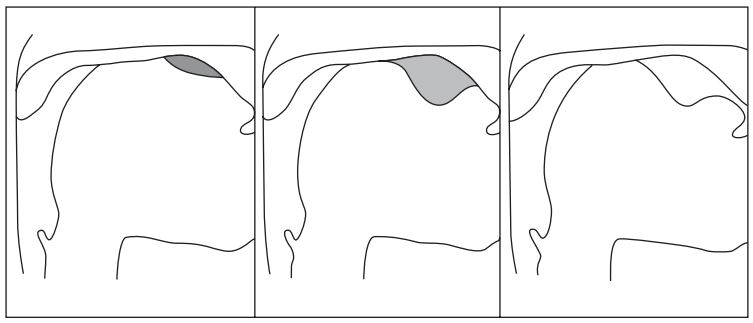


Figure 1 Mechanism for making clicks

Clicks occur in all the Khoisan languages of southern Africa. Five languages of this group, !Xóõ, Ju|'hoan, //Ani, Deti, and Khoekhoe, are included in the sample. The names of the first three of these include click sounds, represented in their spelling by the symbols *!*, *|*, *//* drawn from outside the usual Roman alphabet. Close contact with speakers of Khoisan languages resulted in click sounds being borrowed into languages of the Bantu subgroup of Niger-Congo as speakers of Bantu languages moved south. This borrowing happened independently in two different areas, one toward the east and one toward the west. Zulu (South Africa) represents the eastern and Yeyi (Namibia and Botswana) the western contact area in the language sample. Three languages in East Africa also have clicks. Two of these, Sandawe and Hadza spoken in Tanzania, are sometimes believed to be related to the Khoisan languages further south, but the connection cannot be clearly demonstrated, especially for Hadza. The third East African language with clicks is Dahalo (Cushitic; Kenya). It is usually assumed that the clicks in Dahalo are also the result of borrowing but the source is not known. Because Dahalo also has a "th" type of sound and pharyngeals it falls into a special group on its own.

The occurrence of languages with clicks in East Africa, and especially the relatively far northern location of Dahalo, suggest to many that at one time in the past a much larger area of Africa may have been occupied by people speaking languages that included click sounds, and that these people were displaced or absorbed by later movements of populations coming from the north.

3 Labial-velar plosives

The second class of special consonants are labial-velar plosives. This class of sounds is similar to clicks in an important respect since they also involve two closures, in this case one at the back of the mouth and one at the lips. But in this type the force for the explosive release comes from the air in the lungs as in simple plosives, such as /k/ or /p/. Because of the presence of two closures the phonetic symbols for labial-velar plosives simply combine the two letters used for the separate closures, i.e. /k͡p/ for a voiceless plosive and /g͡b/ for a voiced one. (When it is necessary to distinguish a labial-velar consonant from a sequence of two consonants a tie-bar can be added, e.g. /k͡p/.) Labial-velar plosives occur in 45 (or 8.0 per cent) of the languages in the sample.

Sounds of this type occur in two areas of the world only. The first is an area in West and Central Africa, where /k͡p/, /g͡b/ or both occur in a majority of the languages. Some of these also have a labial-velar nasal, or a prenasalized labial-velar plosive. The greatest number of languages concerned are in the Niger-Congo family, including Temne, Kpelle, Ewe, Dagbani, Kolokuma Ijo, and Yoruba, but some Nilo-Saharan languages such as Ngiti, Lugbara, and Yulu and a few languages in the Chadic subgroup of Afro-Asiatic such as Kotoko also have sounds of this class. The pattern is clearly strongly areal, as it is compact and crosses the boundaries of language families.

The second and quite separate region where labial-velars occur is the eastern end of New Guinea. Two Papuan languages with labial-velars are included in the survey, but there are several others, such as Kâte and Dedua, in the area. Amele (Madang family) has only the voiced plosive /g͡b/, but Yelî Dnye (Yele family; Rossel Island) has several labial-velar consonants (a plain plosive, pre- and postnasalized stops, and a nasal) and in addition sets of consonants which

combine a bilabial closure with an alveolar one or a postalveolar one. It is the only language known to have such a large number of consonants involving combinations of a labial closure with another one.

4 Pharyngeals

The third type of consonants considered in this chapter are those in which the constriction producing the characteristic sound is made in the pharynx. The pharynx is the space above the larynx and below the opening that connects the nasal passage to the mouth (that is, when a person is standing upright). The root of the tongue can be moved backwards towards the back wall of this space to produce a narrowing for a consonant. It can even be moved far enough back that the tongue—or the epiglottis at the base of the tongue—closes against the back wall of the pharynx so that the flow of air from the lungs is blocked and a pharyngeal stop is produced. One or more pharyngeal consonants occur in 24 (or 4.2 per cent) of the languages surveyed. This includes languages with both pharyngeals and the "th" type of sound, one of which (Dahalo) also has clicks. These languages therefore form separate groups in the value list.

Pharyngeal consonants are considerably less concentrated than clicks or labial-velars. They are found in a number of languages in several branches of the Afro-Asiatic family, including Modern Arabic Standard, Tigre, and Soqotri in the Semitic branch, Middle Atlas Berber and Tashlhiyt in the Berber branch, and Somali, Dahalo, and Iraqw in the Cushitic branch. The Nilo-Saharan language Tama is also reported to have a pharyngeal fricative in its inventory. This results in a moderate concentration of languages with pharyngeals in northern and eastern Africa and in adjoining Arabia.

A second area of concentration of the occurrence of pharyngeals is in the Caucasus. Languages of two families in the area, Nakh-Daghestanian (e.g. Ingush, Avar, Rutul, and Hunzib) and North-West Caucasian (e.g. Kabardian and Abkhaz), include pharyngeal consonants. Kabardian is one of the three languages with both pharyngeal consonants and a "th" type of sound. Outside of these two areas of concentration, there are scattered languages in the sample which also include pharyngeal consonants, showing that this class of consonant has clearly arisen independently in areas where contact cannot be a factor. These scattered languages include Nenets (Uralic; Siberia), which is another language also having a "th" sound, some dialects of Atayal (Austronesian; Taiwan), and two languages in the Pacific North-West, Squamish (Salishan; British Columbia) and Nuuchahnulth (Wakashan; British Columbia).

5 "Th" sounds

The final class of consonants to be discussed in this chapter are dental or alveolar nonsibilant fricatives, sounds similar to the "th" sounds of English. These sounds are fricatives made by placing the tip of the tongue behind or just below the upper teeth (when the tip may protrude between the teeth) so that—as with all fricatives—a turbulent flow of air is caused when air is pushed out from the lungs. Crucially, this jet of air is not forced against the teeth as in sibilant sounds such as those usually represented by the letters *s* or *z* (as in English words such as *hiss* and *buzz*) and consequently the sound produced is less loud and lower-pitched than in the *s* or *z* sounds. Dental or alveolar nonsibilant fricatives are written in phonetic notation with the symbols [θ] and [ð] for the voiceless (as in *thing*) and voiced (as in *that*) fricatives respectively.

Dental or alveolar nonsibilant fricatives are just as rare as labial-velar plosives, occurring in just 43 (or 7.6 per cent) of the languages surveyed, but the distribution of these languages is practically worldwide. They are found in languages as varied in location and family affiliation as Modern Greek, Albanian, Spanish, and English (Indo-European), Kabardian (North-West Caucasian), Meadow Mari and Nganasan (Uralic), Burmese and Sgaw Karen (Sino-Tibetan), Lakkia and Yay (Tai-Kadai), Swahili and Moro (Niger-Congo), Dahalo (Afro-Asiatic), Berta and Murle (Nilo-Saharan), Fijian, Yapese, and Drehu (Austronesian), Ngiyambaa (Pama-Nyungan), Rotokas (West Bougainville), Aleut (Eskimo-Aleut), Chipewyan (Athapaskan), Acoma (Keresan), Maricopa (Yuman), Cubeo (Tucanoan), Huastec (Mayan), Mixtec languages and Mezquital Otomí (Oto-Manguean), Amahuaca (Panoan), Tacana (Tacanan), Cochabamba Quechua, and Mapudungun (Araucanian).

One known historical origin of fricatives of this type is the loss of full closure in a dental or alveolar plosive, as happened long ago in the history of English. In some languages a [θ] or [ð] pronunciation may still vary with a plosive pronunciation in different positions in a word or phrase. For example, the spelling *d* in Spanish most often represents [ð] but receives the pronunciation [d] next to certain other consonants or at the beginning of an utterance. Another source of "th" sounds is the loss of the sibilant quality in a fricative of the /s/ or /z/ type, as has occurred in some varieties of Yue Chinese (though not in the Hong Kong Cantonese variety included in this survey). The scattered geographical distribution of dental or alveolar nonsibilant fricatives suggests that the processes that give rise to them are ones that are easily triggered spontaneously. However, the comparative rarity of sounds of this class suggests that these sounds may also be easily lost over a period of time. This may be because of their quietness, which makes them among the least perceptually salient of consonants.

6 Discussion

In contrast to the "th" sounds, the compact geographical distribution of clicks and labial-velar plosives suggests that the evolution or adoption of sounds of these two classes in the sound system of a language is strongly influenced by hearing these sounds in other languages spoken in the same area. Pharyngeals seem to be intermediate in their behaviour, having both a significant tendency to be developed and maintained in particular areal-genealogical groupings of languages and a certain probability of evolving in languages isolated from areas where they are more commonly encountered.

The types of consonants discussed here can all for one reason or another be considered to be among the more complex ones. As noted in Chapter 1, consonants that are evaluated as being more complex in some way in their production are predicted to be more frequently found when the size of the consonant inventory is large, according to what is known as the size principle. This prediction was borne out in the examination of glottalized consonants (see Chapter 7). It is also found with respect to the various types of special consonants discussed in this chapter, as can be seen by comparing the map in this chapter with that in Chapter 1. The relevant distribution of the special consonants is summarized in Table 1.

Table 1 Languages with special consonants by consonant inventory size

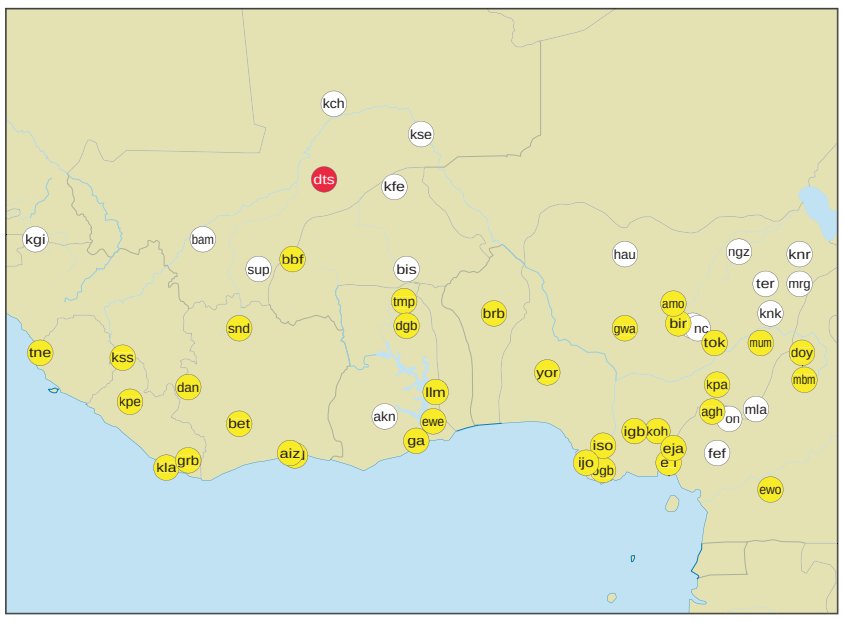
Consonant inventory size class	per cent with any of the special C's
small	8.7%
moderately small	13.1%
average	22.1%
moderately large	27.4%
large	40.7%

The table shows the percentage of languages falling within each consonant inventory size class that have at least one of the special consonant classes among the consonants constituting their inventory. Only a few of the languages with small consonant inventories have any of these consonants, whereas just over two-fifths of those with large inventories include one or more of them. Most importantly, the proportion increases with each increase in overall inventory size. The predictions of the size principle are confirmed for these special consonants in the languages of this sample.

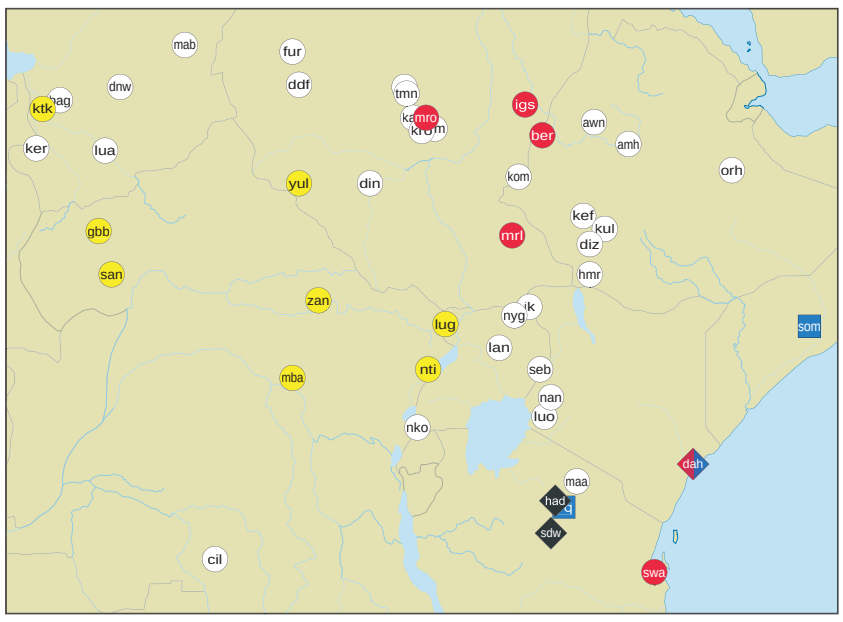
19 Presence of Uncommon Consonants

IAN MADDIESON

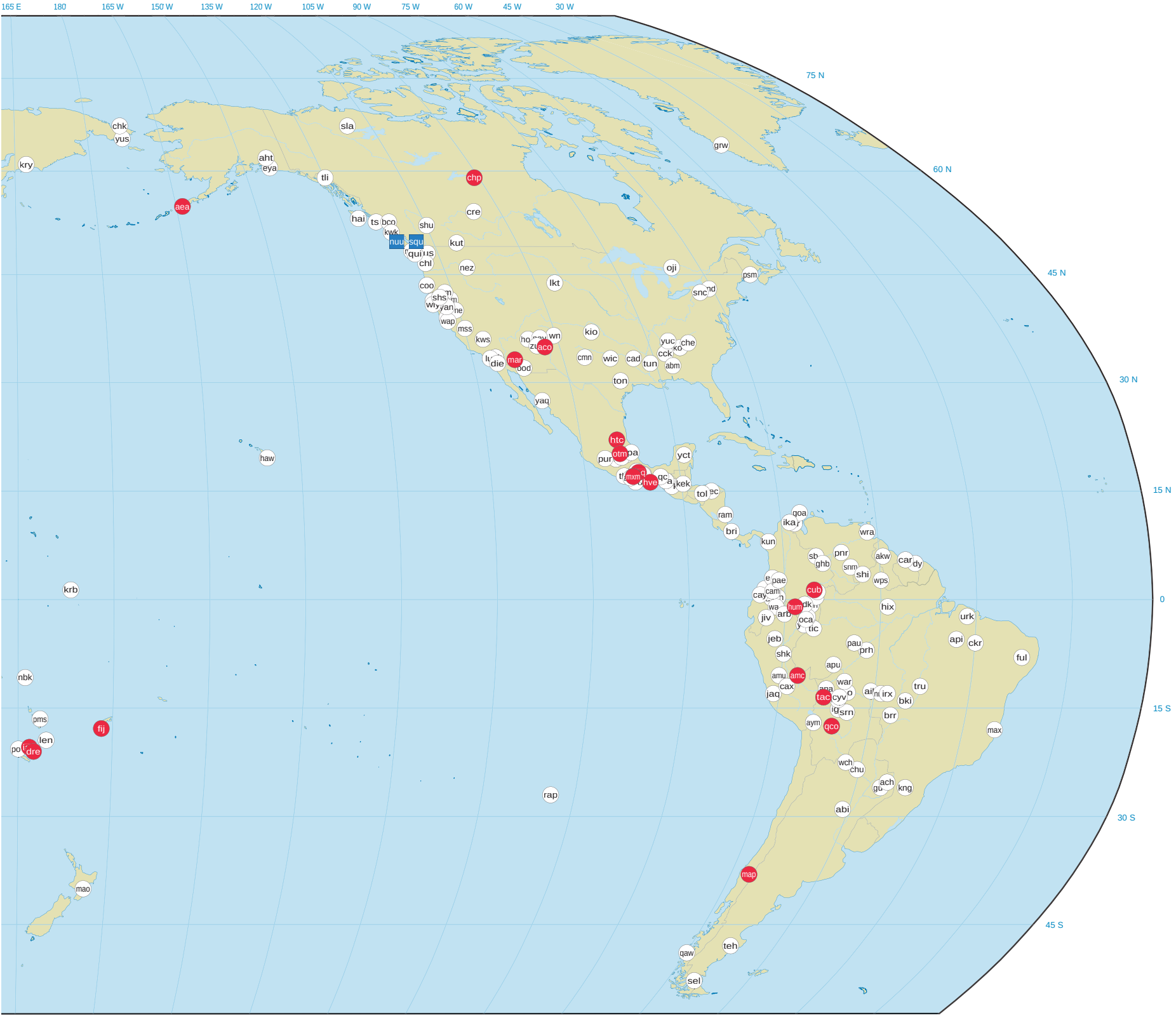
None [448]
Clicks [9]
Labial-velars [45]
Pharyngeals [21]
“Th” sounds [40]
Clicks, pharyngeals, and “th” [1]
Pharyngeals and “th” [2]

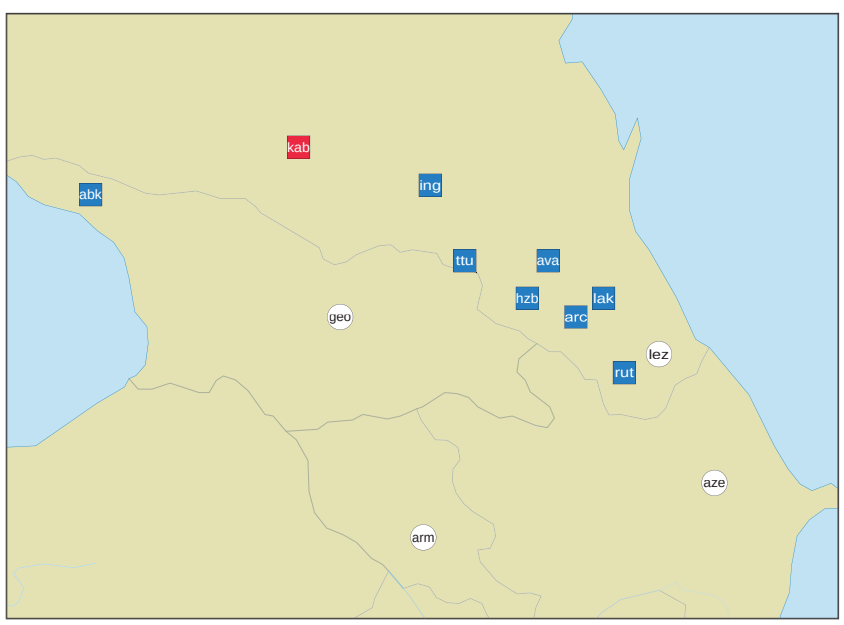


West Africa

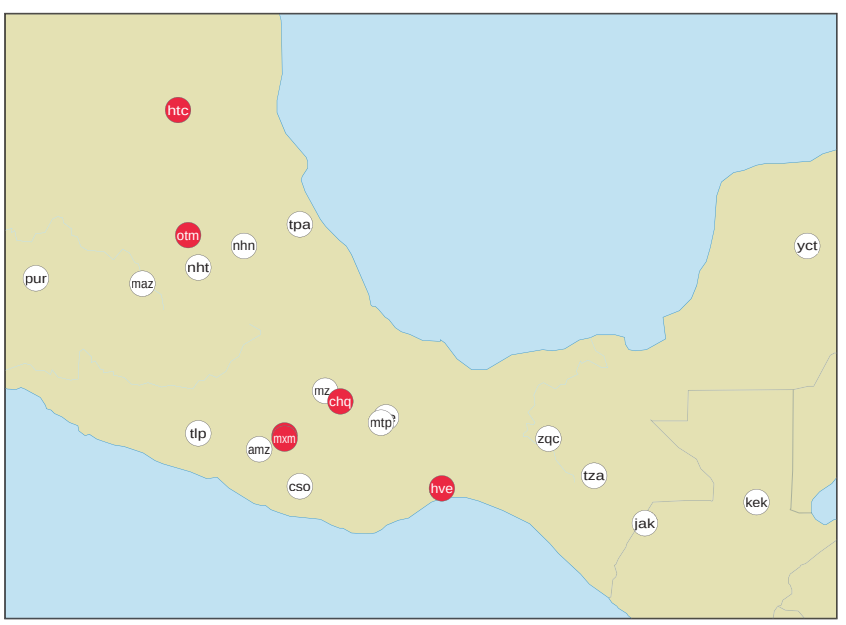


eastern Africa





Caucasus



Mesoamerica

20 Fusion of Selected Inflectional Formatives

BALTHASAR BICKEL AND JOHANNA NICHOLS

Since its beginning in the nineteenth century, morphological typology has postulated a universal scale of less versus more tightly packed word forms. The scale ranges from **isolating** to **agglutinative** to **fusional** to **introflexive**, and is canonically exemplified by Chinese (isolating), Turkish (agglutinative), Latin (fusional), and Modern Standard Arabic (introflexive). Recent research has shown that such a scale conflates many different typological variables and incorrectly assumes that these parameters covary universally (see Plank 1999, Bickel and Nichols 2005, among others). Three prominent variables involved in this are phonological **fusion**, formative **exponence**, and **flexivity** (i.e. allomorphy, inflectional classes). In this chapter we concentrate on fusion. For exponence, see Chapter 21, and for one aspect of flexivity, see Chapter 59 (on possessive classification).

1 Defining fusion types

Fusion refers to the degree to which grammatical markers (called *formatives* in the following) are phonologically connected to a host word or stem. There are three basic values: isolating, concatenative, and nonlinear.

Isolating formatives are full-fledged phonological words of their own. In Fijian, all formatives with more than one mora are isolating. An example is the past tense formative *aa*:

(1) Boumaa Fijian (Dixon 1988: 53)

Au	***aa***	*soli-a*	*a=niu*	*vei*	*ira.*
1SG	PST	give-TR	ART=coconut	to	3PL

'I gave the coconut to them.'

Concatenative formatives are phonologically bound. They need some other host word for their pronunciation and form one single phonological word together with that host. The usual effects of this are that concatenative formatives cannot be individually stressed, and that the combination of formative and host undergoes various phonological adjustments. The past tense marker of Turkish, for example, undergoes vowel harmony and assimilates in consonant voicing to the host stem. Thus, the past tense formative is *-ti* after a stem with unrounded front vowels and a voiceless final consonant (e.g. *git-ti* 'go-past'), *-tı* after a stem with unrounded back vowels and a voiceless final consonant (e.g. *yap-tı* 'do-past'), *-di* after a stem with unrounded front vowels and a voiced final consonant (e.g. *gel-di* 'come-past'), and so on. A subset of concatenative markers is constituted by cliticized words. The Spanish object marker *a*, for example, is syntactically a word (preposition) but phonologically it is a clitic and thus concatenative.

Once the phonological alternations are properly analysed, strings of concatenative formatives can be segmented into clear-cut morphemes. **Nonlinear** formatives are not amenable to this because they are realized not in linear sequence but by direct modification of their host. In our sample, we found two subtypes of nonlinear formatives: **ablaut** and **tonal**. Modern Hebrew illustrates the ablaut type. The past ("perfect") versus future ("imperfect") opposition, for example, is expressed by (i) the choice of a stem template (e.g. *CaCVC* in the past, *CCVC* in the future) and (ii) the choice of agreement affixes (entirely suffixes in the past, mostly prefixes in the future).

(2) Modern Hebrew (Orin Gensler, p.c.)

a. *šamar-ti*	b. *ʔe-šmor*
guard.PST–1SG.PST	1SG.FUT-guard.FUT
'I guarded'	'I will guard'

Neither affix nor stem choice appears to be basic. Tense is not marked in this language by an extractable morpheme but by the complete affix-plus-stem pattern as a whole. For purposes of this survey we call this *ablaut morphology.*

Suprasegmental nonlinear formatives chiefly involve tonal modification. In Kisi (Atlantic; Guinea), most tense-aspect oppositions are expressed by tone, and tone alone:

(3) Kisi (Childs 1995: 220–3)

a. *Ò* *cìmbù.*
   3SG leave.PRES.HABITUAL
   'She (usually) leaves.'

b. *Ò* *cìmbú.*
   3SG leave.PST.PFV
   'She left.'

Here, present habitual is expressed by low tone on the last syllable (3a); past perfective is expressed by high tone (3b).

2 Sampling procedure and feature values

Languages were surveyed for the case and tense-aspect-mood exemplars as defined in the section on "Sampling case and tense formatives" (see box in this chapter). Nearly 90 per cent of the languages in our sample have the same values for case and tense-aspect-mood, and therefore we combined the two values into one overall value:

1. Exclusively concatenative	122
2. Exclusively isolating	16
3. Exclusively tonal	3
4. Tonal/isolating	1
5. Tonal/concatenative	2
6. Ablaut/concatenative	5
7. Isolating/concatenative	13
total	162

Mixed types mean that case and tense-aspect-mood differ from each other, and the type list exhausts what combinations are attested in our sample. In a few instances, however, mixed types refer to languages where there is conflicting evidence for the fusion type of at least one of the formatives. Conflicting evidence is found, for example, in Lakhota (Siouan; North and South Dakota), where the future tense marker *-kta* is part of the same phonological word as the verb stem with regard to morphophonological rules, but not apparently with regard to syllabification (see Russell 1999 for careful discussion). Other instances of conflicting evidence that we found include Beja, Thai, Chamorro, and Gooniyandi.

The colour scheme on the map is set up so as to highlight the presence of some isolating and the presence of some nonlinear formatives.

3 Results and discussion

Most of the languages in our sample (75 per cent) rely exclusively on concatenative morphology for case and tense-aspect-mood. Languages with some isolating or nonlinear formatives are much rarer and have limited areal distribution.

Languages with isolating formatives, or traces of isolating structure in mixed types, are mostly confined to the Sahel Belt of West Africa and to South-East Asia and the Pacific. Apart from this, there are outliers in southern Africa (Khoekhoe), Australia (Gooniyandi), and the Americas (several instances).

In our sample, ablaut morphology is always mixed with concatenative morphology and appears as an African singularity limited to representatives of Afroasiatic (Hebrew, Egyptian Arabic, Middle Atlas Berber, Beja) and the Central Sudanic branch of Nilo-Saharan (Lugbara). Similarly, tonal morphology is also by and large restricted in our sample to African languages (Niger-Congo and Nilotic). An outlier with tonal formatives is Iau (Lake Plains; Papua, Indonesia).

Sampling case and tense formatives

There can be considerable variation in the morphological typology of different formatives in the same language (cf. Plank 1999), especially with regard to fusion (this chapter) and exponence (Chapter 21). The Fijian example in the main text of this chapter illustrates in the same language an isolating tense marker and a concatenative transitivity suffix. In Brahui (North Dravidian; Pakistan; Andronov 1980), case and number are cumulated (i.e. expressed in a single formative) in the nominative, but in the accusative and other cases, number and case are each marked by specialized morphemes (where *-t(ē)* marks the plural; see Table 1).

Table 1 Selected Brahui declension forms (*xal* 'stone')

	SINGULAR	PLURAL
NOMINATIVE	*xal*	*xal-k*
ACCUSATIVE	*xal-ē*	*xal-t-ē*
DATIVE	*xal-ki*	*xal-tē-ki*
ABLATIVE	*xal-ān*	*xal-tē-ān*

This makes it impossible to typologize whole languages for fusion and exponence. In response to this, we sampled individual formatives, one case (or case-like) formative and one tense-aspect-mood (or tense-like) formative. The procedure was as follows (following Bickel and Nichols 2002):

(i) If there is any difference in the morphological type across case formatives, pick the grammatical cases. Within grammatical cases, pick accusative or ergative or agentive (or whatever is chiefly used on A or P arguments). If there is none of these, pick nominative or absolutive (if these are at all marked overtly). If neither the A nor the P argument of transitive clauses is identified as such by overt marking, or if case-marking is restricted to pronouns, assume the language has no "case".

(ii) If there is any difference in the morphological type across tense-aspect-mood formatives, pick tense. Within tenses, pick past (or whatever is chiefly used for simple past time reference); if there is none, pick future; if there is none, pick present. If there is no tense, pick the closest aspect equivalent of past tense as a proxy. If there is no aspect, pick that mood, status, or evidentiality formative that is mostly used for past tense narration. If there is no grammatical marker for any of these notions, assume the language has no "tense-aspect-mood".

(iii) For both case and tense-aspect-mood: if the marking is zero, pick the overtly marked opposite value of the category (e.g. the plural of nominatives, if the singular is zero-marked; or the future tense, if the nonfuture is zero-marked).

(iv) For both case and tense-aspect-mood: if categories differ in their degree of grammaticalization, pick the most nearly grammaticalized one. Pick synthetic tense formatives over periphrastic ones.

Sampling of tense-aspect-mood as defined here was generally straightforward. The most common proxy for past tense was perfective or completive aspect (fourteen languages). In some languages, the proxy was realis status (three languages). In all other languages, tense-aspect-mood morphology was either morphologically homogeneous, or we could identify some dedicated form used for past tense reference.

The sampling procedure for case as defined here mostly revealed ergatives and accusatives. As a result, a language like Brahui (see Table 1 above) will be coded as having a monoexponential case formative even though the nominative apparently cumulates case and number.

For Austronesian languages, we chose the nominative or "topic" form. This form, exemplified here for Tagalog, codes that argument role which the verb is oriented to.

(1) Tagalog (Kroeger 1993: 13)

a. *Bumili* *ang=lalake* *ng=isda* *sa=tindahan.*
PFV.A.buy NOM=man GEN=fish DAT=store
'*The man* bought fish at the store.'

b. *Binili* *ng=lalake* *ang=isda* *sa=tindahan.*
PFV.P.buy GEN=man NOM=fish DAT=store
'The man bought *the fish* at the store.'

c. *Binilhan* *ng=lalake* *ng=isda* *ang=tindahan.*
PFV.D.buy GEN=man GEN=fish NOM=store
'The man bought fish *at the store*.'

The A vs. P vs. D(ative) orientation on the verb specifies the role of the nominative NP (marked by the proclitic *ang=* and given in italics in the translation).

A similar situation is found in Algonquian languages. Here, the NP marked by what is called the "proximative" (zero-marked in opposition to the "obviative") codes that role which the verb is oriented to. Verb orientation is called "direct" versus "inverse" marking by Algonquianists. The following example is from Plains Cree:

(2) Plains Cree (Wolfart 1973: 25)

a. *Sēkih-ē-w* *nāpēw* *atim-wa.*
scare-DIRECT-3 man.PROX dog-OBVIATIVE
'*The man* scares the dog.'

b. *Sēkih-ik* *nāpēw-a* *atim.*
scare-INVERSE[-3] man-OBVIATIVE dog.PROX
'The man scares *the dog*.'

The inverse versus direct orientation specifies the role of the proximative as either A (in 2a) or P (in 2b).

Both Austronesian nominatives and Algonquian proximatives identify argument roles in interaction with verb morphology. Although these types of markers are not traditionally analysed as case, they involve the same basic mechanism as case in languages traditionally assumed to have case: the semantic role expressed by a Latin or Russian nominative depends in part on whether the verb is active or passive (verb orientation).

When there are no markers identifying the roles of A and P arguments, we coded the language as having no case. Thus, in languages such as French, the only (nonspatial) argument role marker is a dative preposition (*à*) used for recipient and goal arguments of ditransitives ('give', 'send', 'tell', etc.). Languages like these were counted as having no case. But if datives or dative-like markers are also used to identify monotransitive objects, they were counted. Marking of monotransitive objects was also counted as case when it is used only on a subclass of objects. An example of this is Turkish, where the accusative is used only with definite objects:

(3) Turkish (Lewis 1967: 35–6)

a. *Mavi* *kumaş-ı* *seç-ti.*
blue material-ACC choose-PST[-3SG]
'She chose the blue material.'

b. *Bir* *mavi* *kumaş* *ist-iyor.*
a blue material want-IMPF[-3SG]
'She wants a blue material.'

Another example is Mandarin, where the formative *bǎ* identifies contextually salient, but not also other, objects:

(4) Mandarin (Li and Thompson 1981: 486)

a. *Tā* *bào-zhe* *zāng* *yīfu.*
3SG hold-DUR dirty clothes
'S/he was holding dirty laundry.'

b. *Tā* *bǎ* *zāng* *yīfu* *bào-zhe.*
3SG OBJ dirty clothes hold-DUR
'S/he was holding the dirty laundry.'

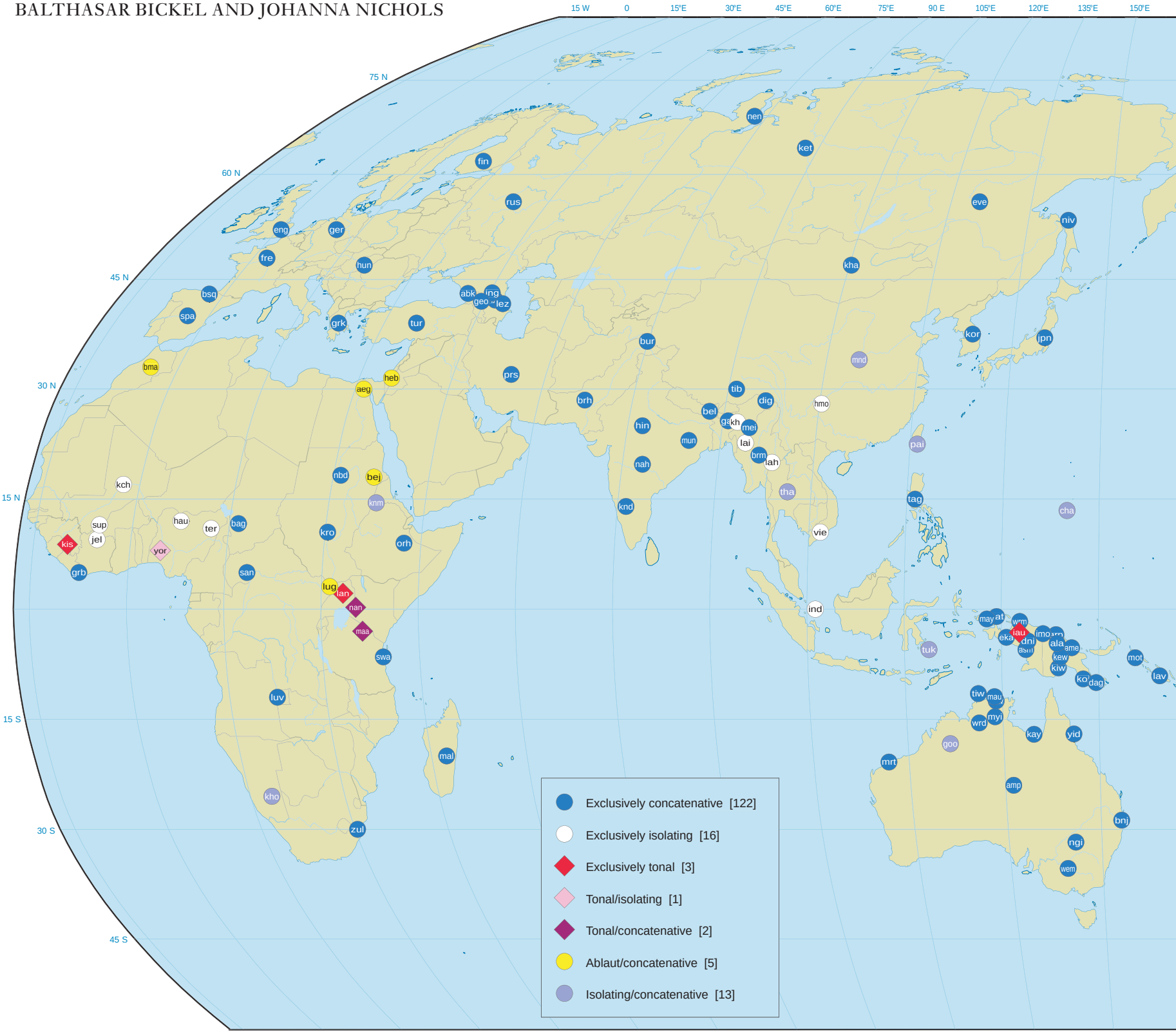
Our notion of case does not differentiate between full-fledged syntactic words (prepositions) and morphological affixes. Hence, the Spanish preposition *a* counts as case. It marks a subset of monotransitive objects.

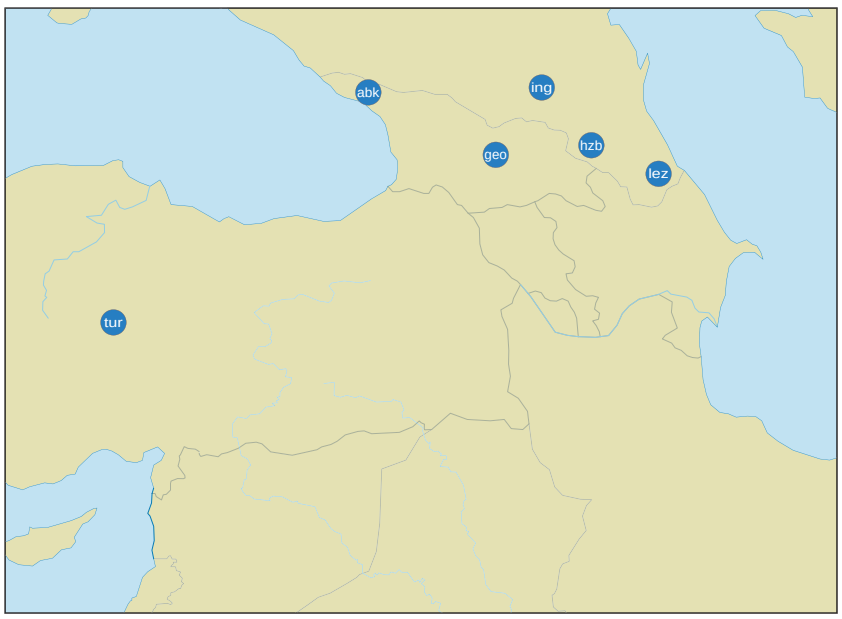
Acknowledgements

This research was supported by US National Science Foundation Grant No. 96-16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P.I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, Dave Peterson, and Suzanne Wilhite for help with data collection.

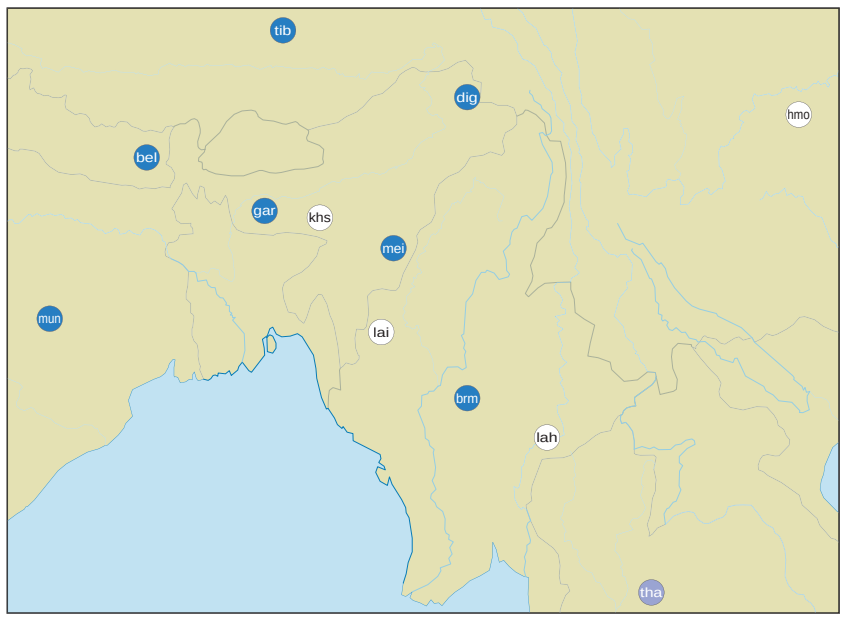
20 Fusion of Selected Inflectional Formatives

BALTHASAR BICKEL AND JOHANNA NICHOLS

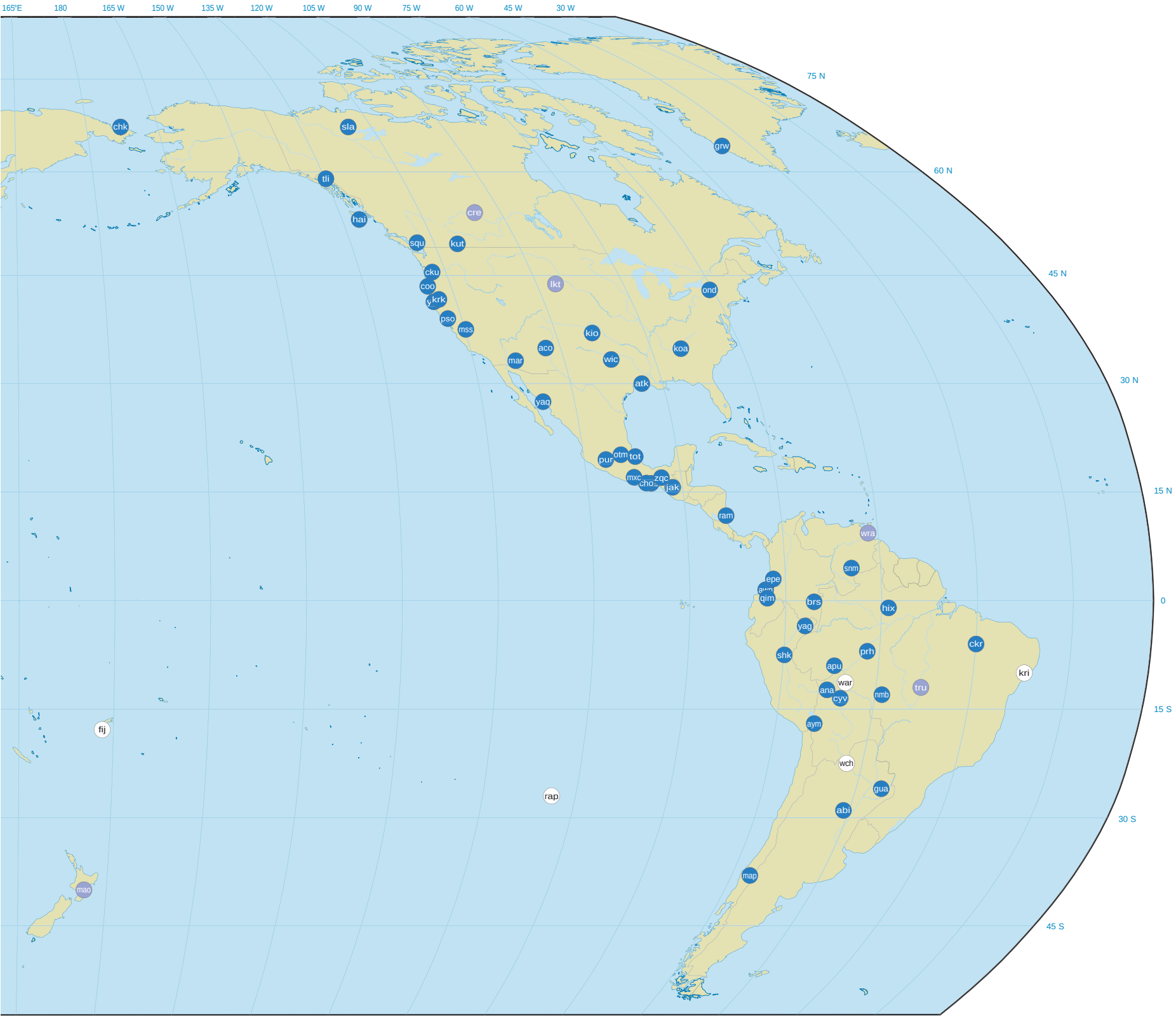


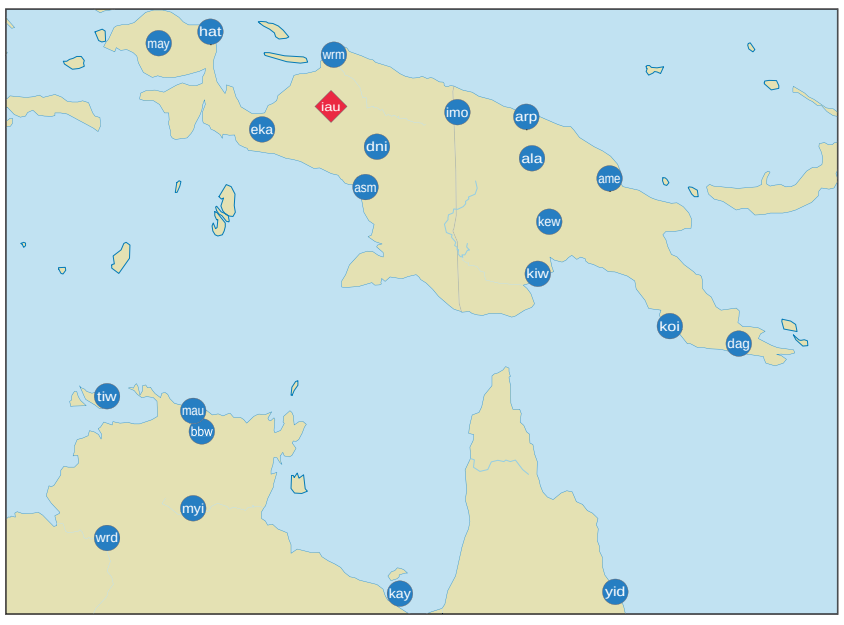


Caucasus and adjacent areas

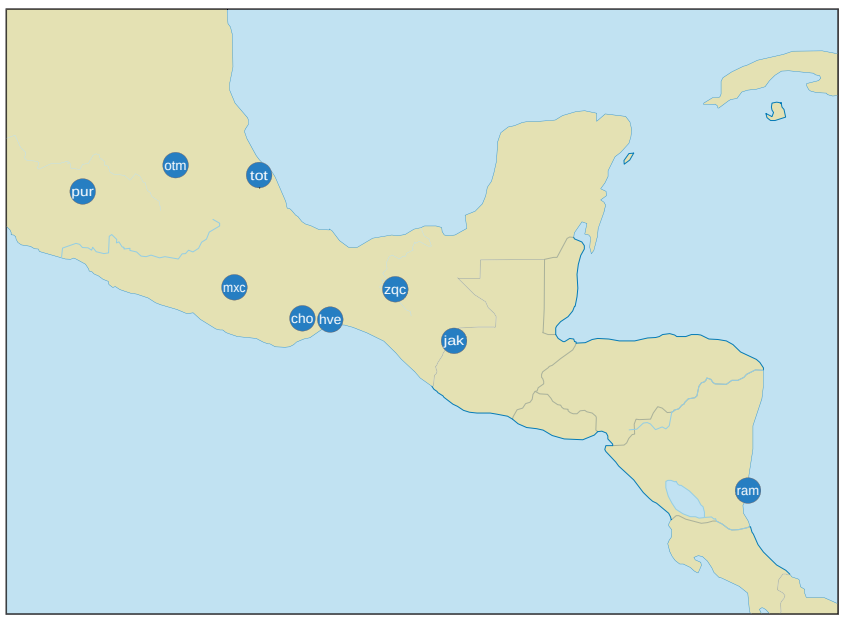


Myanmar and surrounding areas





New Guinea and northern Australia



southern Mexico and Central America

21 Exponence of Selected Inflectional Formatives

BALTHASAR BICKEL AND JOHANNA NICHOLS

Together with fusion and flexivity (i.e. patterns of allomorphy), exponence is one of the parameters of morphological typology. While traditional notions like "agglutinative" and "flexive" conflate these parameters (see Plank 1999, Bickel and Nichols 2005, among others), they are surveyed here separately. For fusion, see Chapter 20; for one dimension of flexivity, see Chapter 59 (on possessive classification).

1 Defining exponence

Exponence refers to the number of categories that cumulate into a single formative. The universal default is to express each category by a dedicated formative. These are **monoexponential** (or separative) formatives. **Polyexponential** (or **cumulative**) formatives, i.e. formatives which simultaneously code more than one category, are much rarer. A well-known example is number and case in many Indo-European languages. Marking of case in these languages involves the same formative as marking of number, and it is impossible to identify separate markers of case and number. Thus, in Russian, genitive singular is *-i* or *-a* and genitive plural *-ej* or *-ov* or *-ø*, and there is no element in these forms that exclusively expresses case or number. When categories are not cumulated into a single formative, each has its own separate marker. This is so in Turkish, where number is always expressed on nouns by means of the suffix *-ler* and for each case there is a separate suffix that can combine with *-ler*, e.g. genitive *-in* (plural *-ler-in*), accusative *-i* (plural *-ler-i*), dative *-e* (plural *-ler-e*), etc.

Exponence is a purely morphological notion; in particular it is independent of the phonological connection between host and formative. Therefore, monoexponential formatives can combine with any fusion type (see Chapter 20 for the definition of fusion types). Thus, the Fijian past tense formative *aa* (see example (1) in Chapter 20) is monoexponential and isolating. The Turkish past tense formative *-ti* (and its alternants) is monoexponential and concatenative. The Kisi high tone past perfective formative is monoexponential and nonlinear.

2 Sampling procedure and feature values

We sampled, for each language, one exemplar of a case and one exemplar of a tense-aspect-mood formative, following the procedure set out in the section on "Sampling case and tense formatives" (see the box in Chapter 20). We were interested in combinations of very different categories in one formative, and not for instance in combinations of some tense value with some aspect value. Thus, for coding purposes we treated tense, aspect, mood, status, and evidentiality as one category, and checked whether or not this cumulates with such categories as agreement, voice, or negation. Many such category combinations are possible, but in our sample, we found only those shown on the maps.

2.1 Case exponence. Case exponence has a different distribution than tense-aspect-mood exponence, but in both instances, polyexponence is rare. The main map displays case exponence, with the following feature values:

1. Monoexponential case	69
2. Case+number	8
3. Case+referentiality	6
4. Case+TAM (tense-aspect-mood)	2
5. No case	75
total	160

Case+number coexponence is illustrated by the Russian example discussed above. By coexponence of case with referential values (**case+referentiality**) we refer to case markers that specify their host NP as topics or as having specific or definite reference. In our sample, this characterizes the Austronesian nominatives (also called "focus" markers or "triggers"), and the Algonquian and Kutenai proximatives. Each of these case markers not only marks a specific grammatical relation but also signals that the NP so marked is topical and definite. For illustration and discussion, see examples (1) and (2) in the explanation of case sampling (see the box in Chapter 20).

While referential values also play a role in differential object (or differential subject) marking (Bossong 1985, 1998), we did not treat such marking as instantiating case+referentiality coexponence. In systems of differential marking, definiteness emerges only secondarily as the result of using or not using a certain case marker in response to the situational salience or referential activation of the object, e.g. using or not using the accusative in Turkish or the *bǎ* particle in Mandarin in response to situational salience (see examples (3) and (4) in the explanation of case sampling (cf. the box in Chapter 20)). Case markers with a true referentiality coexponent are different in that the use of these markers is obligatory: in Tagalog or Plains Cree, one of the NPs in a clause MUST be in the nominative or proximative, respectively. What speakers can choose is on which argument they place the marker, but they cannot choose to simply drop the case-marking altogether. A borderline example of case+referentiality coexponence is Evenki. This language has two accusative forms, one for indefinite objects (2a) and an unmarked one for definite or indefinite objects (2b):

(2) Evenki (Nedjalkov 1997: 193)

a. *Beje-l* *mo:ka-r-e* *genne:-vki.*
man-PL[-NOM] stick-PL-INDEF.ACC bring-HABITUAL.PTCP
'The men usually bring firewood.'

b. *Tar* *asi* *kniga-va* *tang-d'ara-n.*
that woman[-NOM] book-ACC read-PRES-3SG
'That woman is reading a/the book.'

This is like true case+referentiality coexponence because case marking is obligatory with objects, but the functional distribution of the two case markers approximates the differential object-marking pattern.

Case+TAM (tense-aspect-mood) coexponence is an extremely rare phenomenon. In our sample (and to our knowledge in general) there are only two instances. One is Kayardild (Tangkic; Queensland, Australia):

(3) Kayardild (Evans 1995: 404)

a. *Ngada* *kurri-nangku* *mala-y.*
1SG NOM see-NEG.POTENTIAL sea-LOC.ACTUAL
'I could not see the sea.'

b. *Ngada* *kurri-nangku* *mala-wu.*
1SG NOM see-NEG.POTENTIAL sea-PROPRIETIVE.FUT
'I won't (be able to) see the sea.'

The past and future tense values are exclusively established here by the choice of case markers on the object (locative for actual, instantiated events, proprietive for future events). The other instance is Lugbara (Central Sudanic; Uganda). Here, the difference between present and perfect tenses is indicated (at least in some environments) exclusively by the presence versus absence of the subject case marker. (The case marker also cumulates the expression of number.)

(4) Lugbara (Crazzolara 1960: 80–3)

a. *Ɔzɔ́ɔ́=nì* *'dì ~ 'di.*
rain=NOM.SG.INCOMPLETE rain~EMPH
'It is raining.'

b. *Àmbó* *mu* *'dálé.*
bigman go there
'The Bigman has gone there.'

This is different from TAM-based subject-marking splits (split-ergative, split-nominative) because in (4), TAM values are not independently indexed in the clause by the verb form or some particle.

2.2 Tense-aspect-mood (TAM) exponence. The inset map displays exponence of tense-aspect-mood in the same set of languages:

1.	Monoexponential TAM	127
2.	TAM+agreement	19
3.	TAM+agreement+diathesis	4
4.	TAM+agreement+construct	1
5.	TAM+polarity	5
6.	No TAM	4
	total	160

Values for Map 21A Exponence of Tense-Aspect-Mood Inflection

TAM+agreement coexponence is well known from Indo-European and Semitic languages, both ancient and modern. It is also common in Papuan languages, especially in those classified as part of the Trans-New Guinea Phylum (Foley 1986: 137). In Indo-European and in some Papuan languages, the same formatives expressing TAM+agreement also code **diathesis** (e.g. active versus passive versus middle in Modern Greek, active versus middle in Kewa). In one instance, the TAM+agreement exponence pattern also extends to what we call **construct** marking. Construct marking is the overt registration on the head of the presence of a dependent (or a particular kind of dependent, or a dependent in a particular position). This phenomenon is best known from Semitic NPs, where construct forms register the overt presence of a dependent possessor. With verbs, the construct form registers the overt presence of particular arguments. It is chiefly found (but variously labelled) in languages of Africa and the Pacific. The only instance in our sample where verbal construct marking is combined with a TAM formative is Lango (Nilotic; Uganda): there is one perfective marker used after the pronominal subject *én* 'he, she, it' or with relativized subjects (5a), and another perfective marker used elsewhere (5b):

(5) Lango (Noonan 1992: 137)

a. *én* *òcámò*
s/he 3SG.eat.PFV.CONSTRUCT
'He ate it.'

b. *òcàmò*
3SG.eat.PFV
'He ate it.'

The formatives differ from each other by the tonal pattern; however, the tonal pattern does not exclusively mark the construct versus nonconstruct difference but it also encodes (part of) the agreement and aspect features.

TAM+polarity coexponence, finally, is found in languages where the expression of negation in the verb is so tightly interwoven with the coding of TAM values that no morphemes can be extracted for either category.

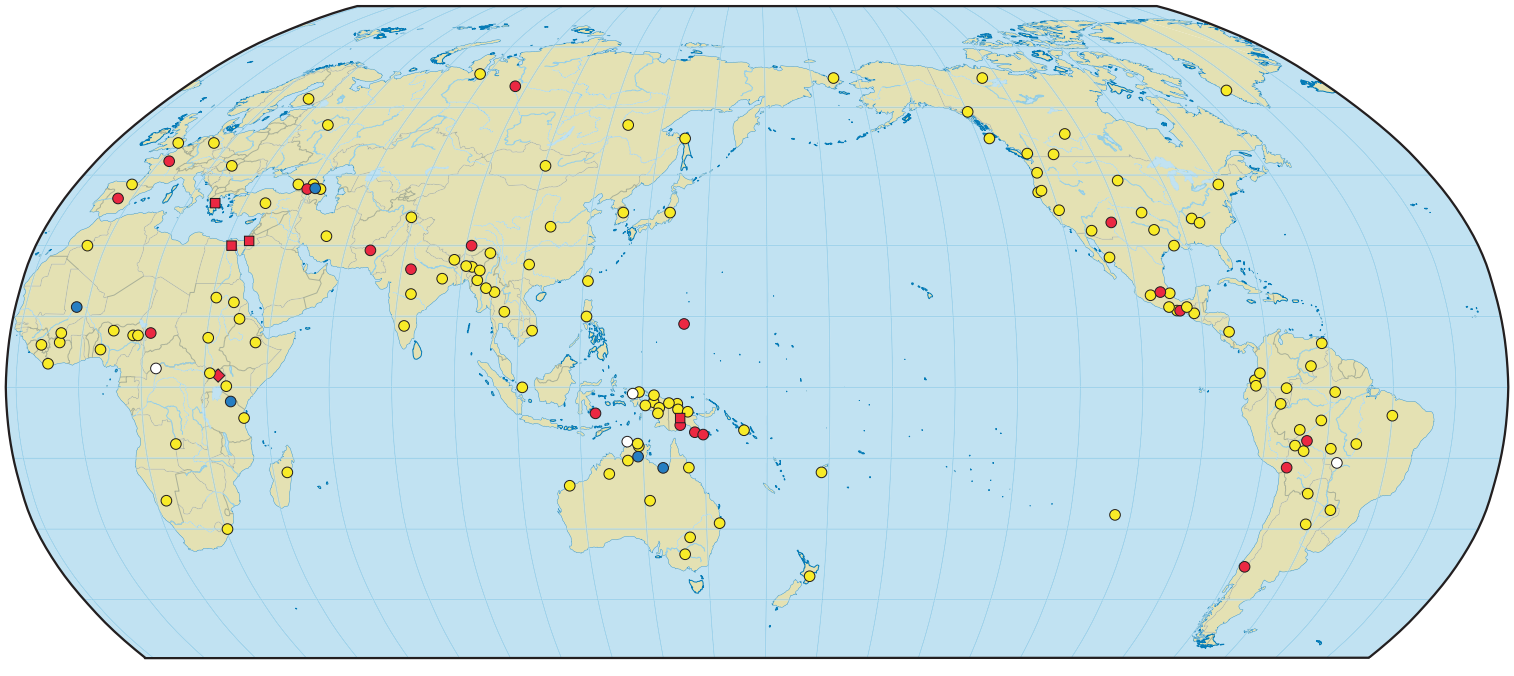
3 Results and discussion

All instances of polyexponence are singularities, limited to single languages or single families. Polyexponence is extremely resistant to areal spread. In turn, some instances are well attested throughout many branches of their families. This suggests great genealogical stability, and current distributions across individual families perhaps point to historical relationships beyond the reach of the comparative method.

An example of this is coexponence of case with number. This pattern is a stable feature of archaic Indo-European case systems, represented here by German, Greek, and Russian. It is also found in several branches of Uralic (here, in Finnish and Nenets). Apart from this, case+number coexponence shows up in Eskimo, Chukchi (Siberia), Yaqui (Uto-Aztecan; northern Mexico), and Lugbara (Central Sudanic; Uganda). However, on closer inspection, in most of these languages the phenomenon is different from the one in Uralic and Indo-European (e.g. limited to individual cases or subsets of nouns as in Chukchi, or resulting from total case-neutralization in the plural as in Yaqui). This suggests that the distribution in Indo-European and Uralic perhaps reflects a singularity inherited from a common protolanguage.

Acknowledgements

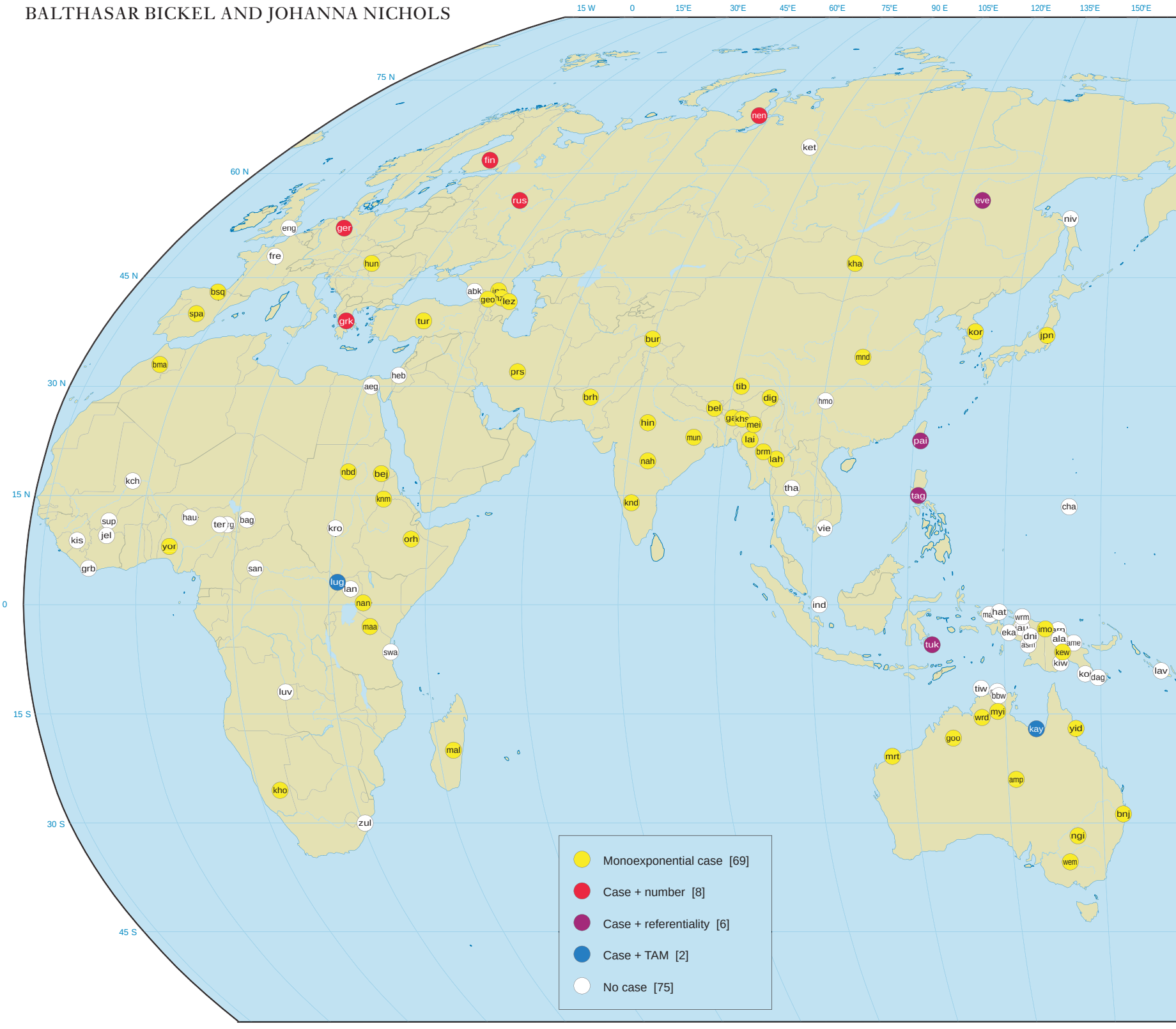
This research was supported by US National Science Foundation Grant No. 96–16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P.I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, Dave Peterson, and Suzanne Wilhite for help with data collection.

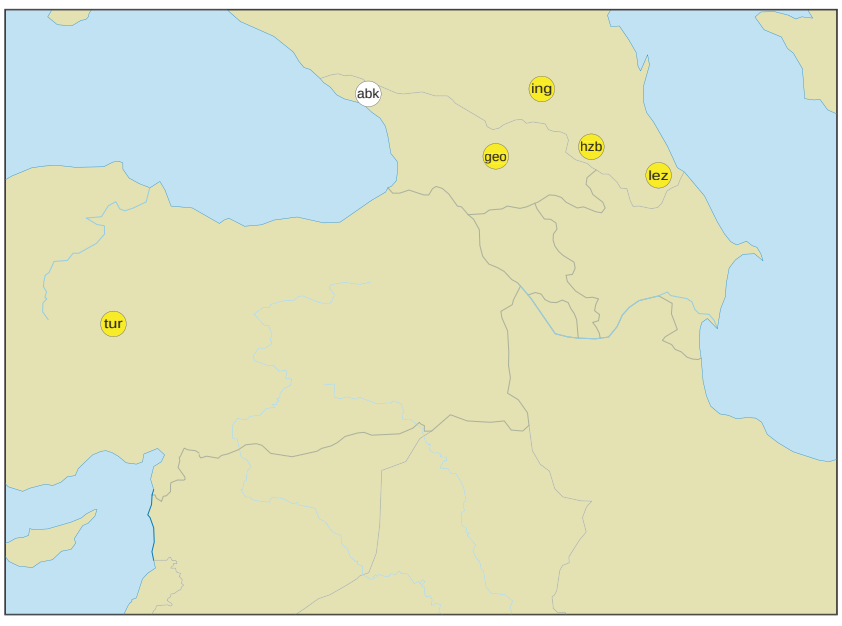


Map 21A Exponence of TAM (tense-aspect-mood) Inflection

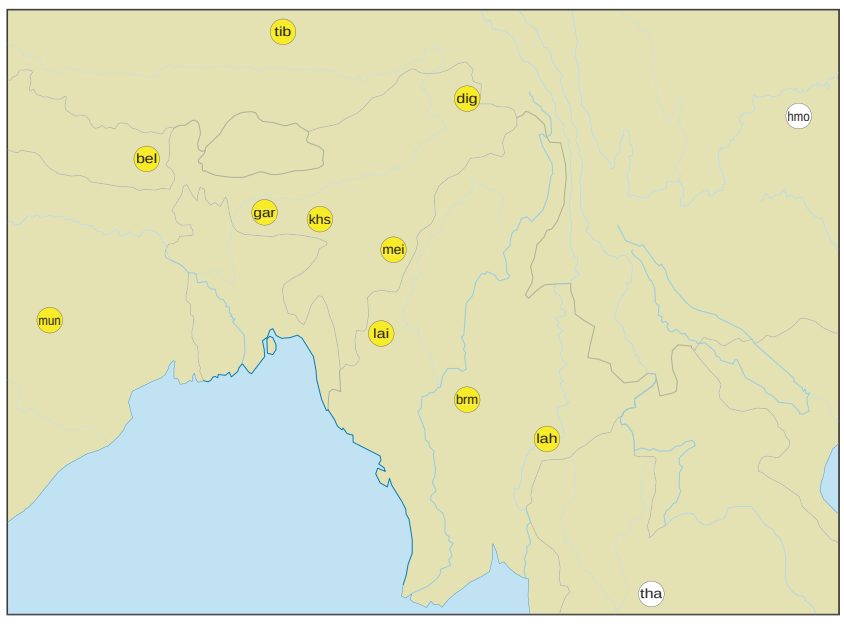
21 Exponence of Selected Inflectional Formatives

BALTHASAR BICKEL AND JOHANNA NICHOLS

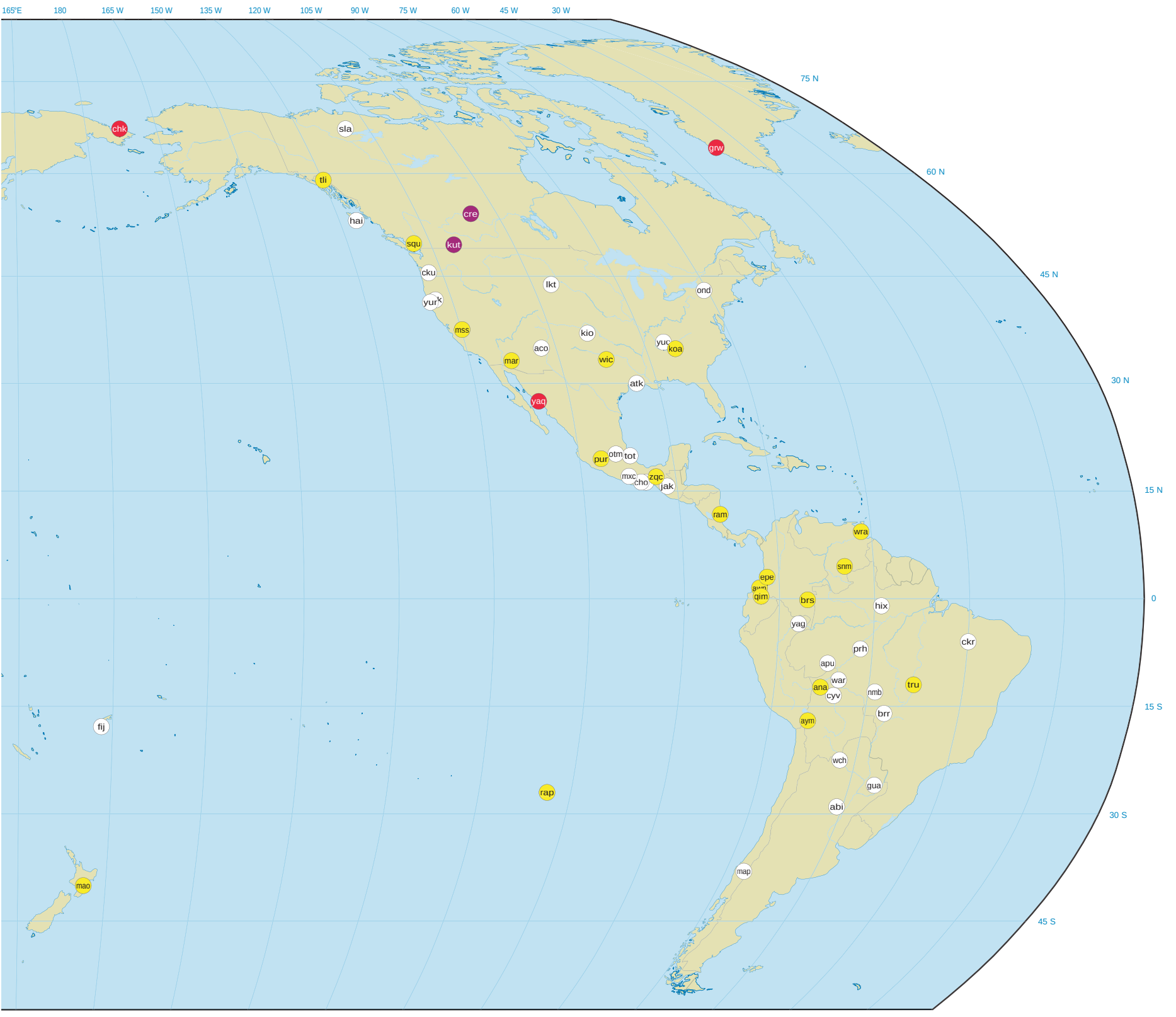


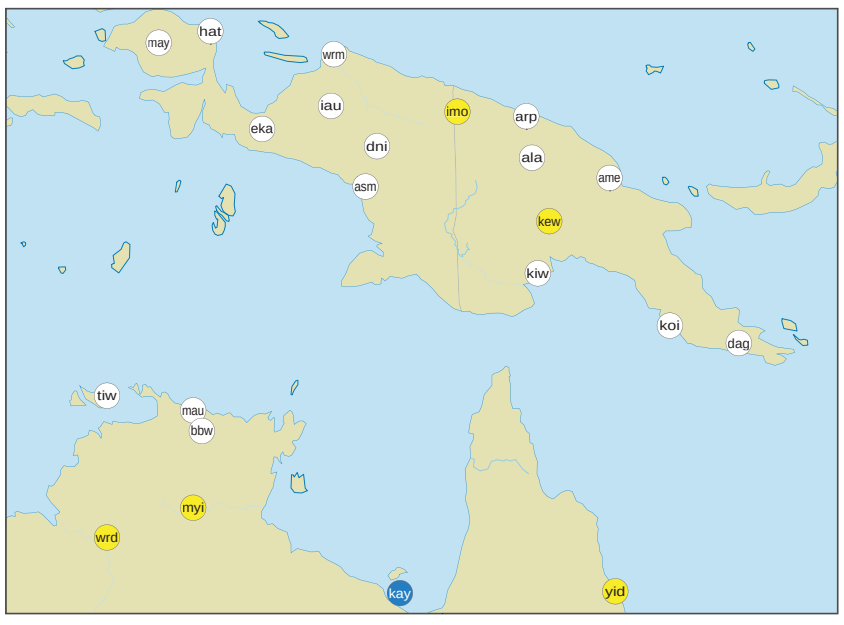


Caucasus and adjacent areas

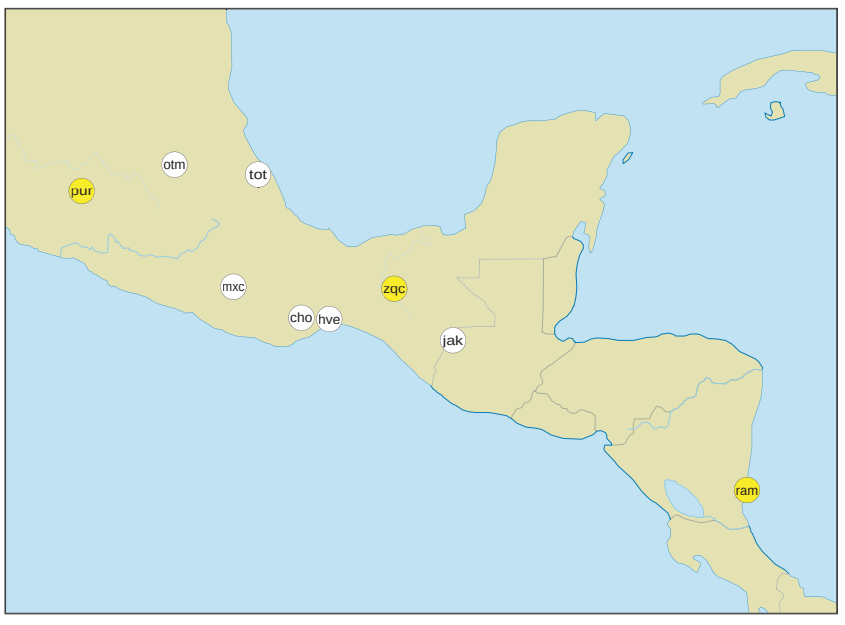


Myanmar and surrounding areas





New Guinea and northern Australia



southern Mexico and Central America

22 Inflectional Synthesis of the Verb

BALTHASAR BICKEL AND JOHANNA NICHOLS

1 Defining synthesis

Grammatical categories like tense, voice, or agreement can be expressed either by individual words or by affixes attached to some other word (or the stem of a word). If a word combines with affixes, the resulting construction is said to be **synthetic**; if not, it is said to be **analytic**. For example, in the past tense, the English verb combines with an affix expressing tense (cf. *she painted*, with suffix *-ed*). Forms like *painted* are called synthetic. In the future tense, by contrast, the verb does not combine with an affix; the future is instead expressed by a separate word (*will*, as in *she will paint*). Expressions like *will paint* are called analytic.

In general, a string of elements makes up one synthetic word when these elements stand in a morphological, and not in a syntactic, relation to each other. For example, *paint* and *-ed* stand in a morphological relation to each other, and they therefore make up one word; *will* and *paint*, by contrast, stand in a syntactic relation to each other, and they are therefore two words. Morphological relations are generally more constrained; they imply fixed order, and the parts involved are unable to appear on their own (e.g. **ed-paint, *ed!*) and cannot be focused individually (apart from metalinguistic correction as in *I said she paintED*). Syntactic relations, by contrast, generally allow (some) re-ordering (*will she paint?*), the parts involved can appear on their own (*she will*) and can be focused (*she WILL paint!*). Moreover, they may involve case government and/or agreement.

Synthesis is independent of phonological (or prosodic) **fusion** (on which see Chapter 20): words can phonologically cliticize to other words, but this is not synthesis. The English auxiliary word *will*, for example, typically cliticizes to preceding NPs (cf. *she'll paint*), but the relationship between *'ll*, *she*, and *paint* is still syntactic, and from the point of view of syntax, there are still the same three words as in the noncliticized version (*she will paint*).

Likewise, what is a single synthetic word from the point of view of syntax can be composed of several distinct phonological words, but that does not make it any less synthetic. In Hakha Lai (Tibeto-Burman; Myanmar; Ken VanBik, p.c.), for example, the future tense marker *làay* synthesizes with verbs: unlike an individual syntactic word, *làay* cannot be re-ordered and it cannot be used alone, without a verb. However, *làay* is still a separate phonological word: like all other phonological words, and unlike clitics in Lai, it containes two moras and bears a tone. Thus a string like *a-nii làay* [3SG-laugh FUT] 's/he will laugh' is one single synthetic word though prosodically two words. (This word form also includes the agreement prefix *a-*.) (The difference between phonological fusion and synthesis is further discussed in Bickel 2003; Bickel and Nichols 2005.)

For phonologically separate particles like the Hakha Lai future tense marker, available descriptions sometimes make it difficult to establish whether or not these particles synthesize with another word. In our survey we made the default assumption that phonologically separate particles are also syntactically separate, i.e. that they do not synthesize. But when these particles manifest clear morphological dependency, e.g. when they cannot be used alone, or when their position is rigidly adjacent to words and this rigid placement contrasts with relatively free word order, then we took this as evidence for synthesis. The contrast between synthesizing and non-synthesizing particles can be illustrated by negation particles in mainland South-East Asia. In Hmong Njua (as well as in Burmese, Sgaw Karen, Hakha Lai, and presumably Thai), negation particles are strictly adjacent to and dependent on verbs (Harriehausen 1990). But in Mandarin (and similarly in Vietnamese), placement of the negation particle *bu* is determined by syntactic principles of phrase structure and by semantic principles of scope. Syntactically, *bu* must be adjacent to the verb phrase (but not necessarily to the verb). Semantically, *bu* precedes whatever is negated. Compare the following examples:

(1) Mandarin (Li and Thompson 1981: 420)

a. *Tā tiāntiān bu xǐzǎo.*
   3SG daily NEG bathe
   'Every day she does not bathe.'

b. *Tā bu tiāntiān xǐzǎo.*
   3SG NEG daily bathe
   'She does not bathe every day.'

In (1a) *bu* negates the bathing (and the time specification is presupposed); in (1b) *bu* negates the time specification (and the bathing is presupposed). In both cases, *bu* is adjacent to a VP (*xǐzǎo* 'bathe' in 1a, *tiāntān xǐzǎo* 'bathe daily' in 1b). Sentences like (1b) are ungrammatical in Hmong Njua (Harriehausen 1990: 173) because the negation particle must be adjacent to the verb.

2 Surveying inflectional synthesis of the verb

Our survey concentrates on the synthesis of inflectional categories with verbs. By inflectional category we understand any grammatical category whose presence or shape is (at least in part) a regular response to the grammatical environment. The prime candidates for this are categories like agreement, tense-aspect-mood, evidentials (or miratives), status (realis, irrealis, etc.), polarity (negation), illocution (interrogative, declarative, imperative), and voice (including Austronesian-style verb orientation). Often, these categories are sensitive to the syntactic environment (e.g. argument NPs in the case of agreement, sequence of tense rules in the case of tense, cross-clausal anaphora in the case of voice, etc.). But often, the grammatical sensitivity is more narrowly morphological: different evidential or tense forms, or negative versus affirmative forms, may imply different paradigms, or combine with different sets of aspect forms, or voices, etc. (see Bickel and Nichols 2005 for discussion).

While information about morphological sensitivity can usually be gleaned from basic descriptions, data on syntactic sensitivity is often unavailable. We made the default assumption that unless there is positive evidence for any kind of grammatical sensitivity of the category, it is not inflectional but derivational.

Apart from the better-known and common inflectional categories, the following categories proved to have verbal inflectional reflexes in at least one language: nominalizers, connectives or switch-reference markers (as in Belhare, Kiowa, Fijian, Daga, Maricopa), inverse marking (as in Plains Cree, Mapudungun, or Chukchi) or Kartvelian-style "version", honorificity (as in Japanese or Korean), pluractionals and other quantificational categories (multiple argument or multiple action, as in Wichita and Koasati, or repetition marking, as in Karok), verb focus or emphasis (as in Maricopa, Pirahã, or Imonda), transitivity markers (as in Fijian, Plains Cree, Krongo, or Hakha Lai), reciprocal affixes (triggered by agreement with free reciprocal pronouns, as in Chamorro), construct marking (indicating the presence of a dependent NP, as in Hausa, Lango, and Supyire), object classifiers (inflectional if interacting with agreement, as in Imonda), nonspecific reference-marking (in Koasati), scope (delimiting the scope of other categories, as in Mezquital Otomí), deixis (judged inflectional in, for example, Lango because it interacts with agreement paradigm rules), and motion (judged inflectional in Yagua and Imbabura Quechua).

In addition, causatives were judged inflectional in some languages, where these categories are regular responses to specific contexts (e.g. a response to specific types of switch-reference patterns as in Ingush) or where they are worked into the same paradigms as regular voice values (e.g. in Chamorro, Fijian, or Malagasy).

3 Measuring synthesis

Within the same language, verbs can be used with more or less synthesis: the English past, for example, is more synthetic than the future. For surveying purposes, we looked for the **maximally inflected verb form**, i.e. the one form that is most synthetic, and determined its **category-per-word value** ("cpw value"). In English, the maximally inflected verb form expresses two categories: agreement (in the present: *-s*) and tense (past: *-ed*). The English verb therefore has a synthesis degree of 2 cpw (=categories per word). At one extreme in the sample is Vietnamese, where we found no evidence for any synthetic inflectional category in the verb. The maximally inflected verb in Vietnamese therefore has 0 cpw. At the other extreme are languages like Koasati, whose inflected verb forms can include up to 13 cpw.

In some languages, the same or a similar category can be expressed at various places in the verb. When these places were sufficiently distinct, we counted the category twice (or more). An example of this is number- or pluractional-marking on the stem in addition to affixal number agreement. Another example is certain kinds of aspect in addition to other kinds of aspect-marking.

On the other hand, when two (or more) semantically related categories cumulated into one single inflectional slot or morpheme, we counted this as one category (see Chapter 21 on exponence). In particular, unless tense, aspect, and mood were clearly distributed in distinct positions, we counted these as one category (thereby avoiding difficult decisions as to whether something is aspect or tense or mood). When categories were not related, however, we treated them as distinct. Hence, cumulated exponence of agreement and tense, as in German and other Indo-European languages, was coded as two categories. All these issues can be illustrated by an example from Kewa (Engan; Papua New Guinea):

(2) Kewa (Franklin 1971: 49–50)

a. *íra-b-e*
cook-INCEPTIVE-1.SG.PERF.EGOCENTRIC
'I have begun cooking it'

b. *íra-paa-ru*
cook-COMPLETIVE-1.SG.PST.ALTROCENTRIC
'I finished cooking it (for someone)'

c. *íra-wa-de*
cook-1.SG.PST.EGOCENTRIC-PUNCTILIAR
'I cooked it'

In such a case, we disregard the issue of whether the perfect is more a tense or an aspect (or neither) and simply count this as one category ("tense-aspect"). This category cumulates with agreement and diathesis-marking (called altrocentric versus egocentric benefaction, but akin to active versus middle voice). Since agreement and diathesis are not related to tense-aspect, we count them each as one category. In addition to the perfect/past marking, there is an inceptive/completive/punctiliar distinction. Since this is marked in a clearly separate position and shows different behaviour from the perfect/past distinction, we count this as an additional, separate aspect category. The Kewa forms thus combine the four categories (aspect, tense-aspect, agreement, diathesis) shown here (plus polarity and evidentiality not shown here).

Agreement was always counted as one category per role (subject agreement, object agreement, etc.), no matter how many features are affected by agreement. Thus, agreement in gender was counted the same way as agreement in person, number, gender, and honorificity, etc. When agreement cumulates with another, non-agreeing, category (e.g. tense-aspect-mood, as in classical Indo-European), we counted both categories individually.

4 Results and discussion

Given that we made default assumptions disfavouring synthesis (as noted above), the cpw values represented on the map are possibly judged lower in less well-described grammars than in the better-known languages. However, our survey showed that there is no correlation between the obtained cpw value and the quality of description, so we are confident that the obtained cpw values are a sufficiently reliable approximation to the worldwide distribution of different degrees of synthesis.

But individual cpw values are in many cases valid only within a confidence interval of around ±1. In response to this, the map depicts cpw values as a scale in spectral colors:

1. 0–1 category per word	5
2. 2–3 categories per word	24
3. 4–5 categories per word	52
4. 6–7 categories per word	31
5. 8–9 categories per word	24
6. 10–11 categories per word	7
7. 12–13 categories per word	2
total	145

From a worldwide perspective, the most common cpw values are between four and eight (acounting for two thirds of the sample) and the universal mode is four (17.4 per cent). Especially outside Africa and Australia, the distribution is geographically very uneven. In particular, Eurasia is dominated by low-synthesis languages and the Americas (especially North America) by high-synthesis languages.

Exceptions to the low synthesis average in Eurasia are found on the Pacific Rim and in the Himalayan and Caucasian mountain zones. Without these three regions, Eurasia is relatively homogeneous and has low synthesis degrees. The values in the Caucasus, by contrast, include values of up to 11 cpw; in the Himalayas and on the Pacific Rim, of up to 8 cpw. These three geographical zones are **typological enclaves** (see Bickel and Nichols 2003 for theory and statistical analysis of this phenomenon). They are relic zones with deviations from the surrounding typological profile and these deviations can be very heterogeneous, as is the case, for example, in the Caucasus.

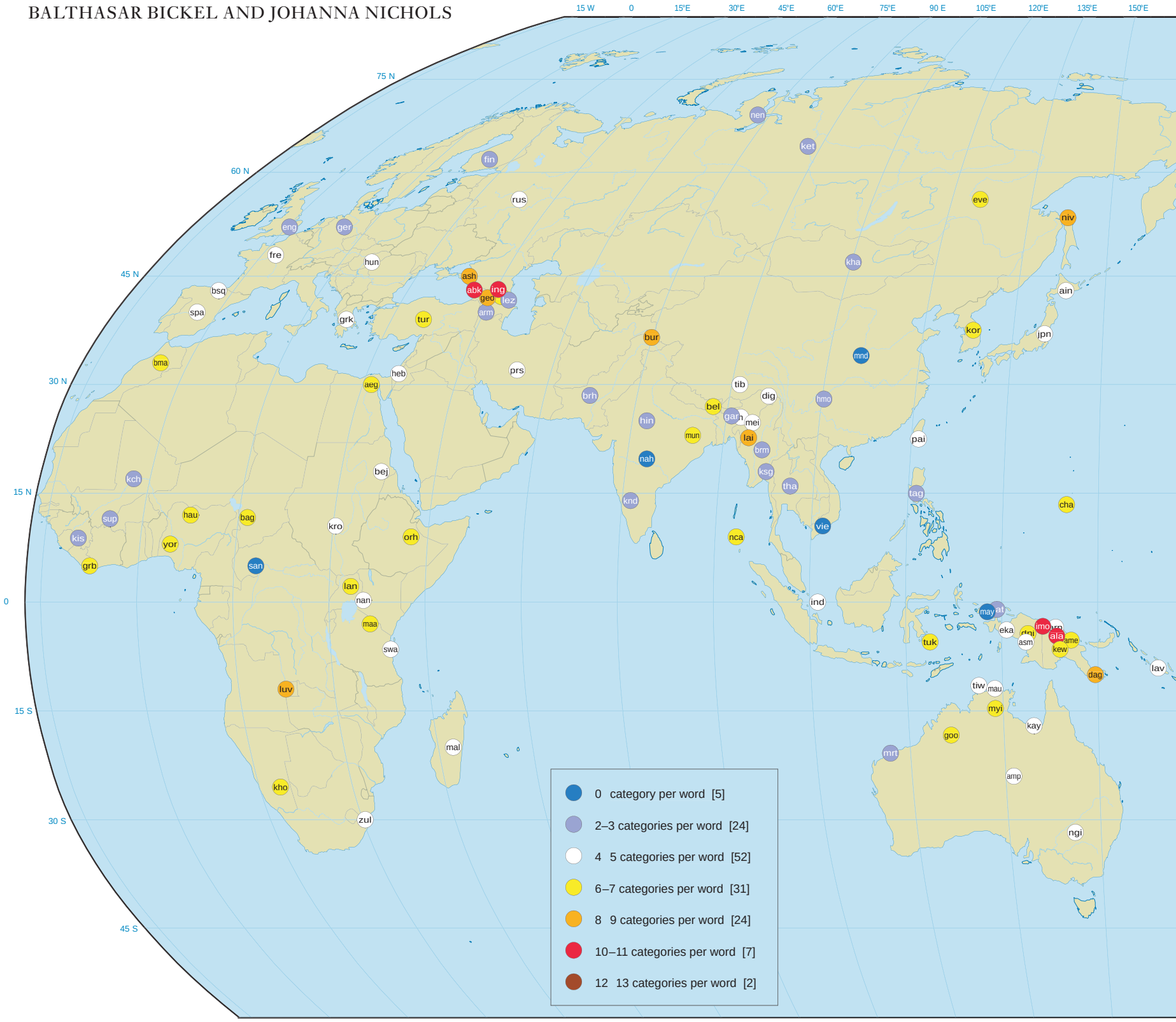
There is an outlier at the fringes of each Eurasian enclave: Turkish near the Caucasus and Lai near the southeastern foothills of the Greater Himalayas. The high synthesis of Turkish reflects a common trend throughout the Turkic language family (deep into Siberia). It seems that this family has not adjusted its typological profile to the Eurasian standard (perhaps because that profile was established before Turkic speakers started to move westward in the first millenium CE, and not enough time has since elapsed for Turkic to lose high synthesis degrees and therefore assimilate to the standard). Lai apparently reflects the typological profile of Tibeto-Burman before the great spread of this family into South-East Asia. Outside the Kuki-Chin and a few other branches of Tibeto-Burman, this profile survives only in the archaic Kiranti group that is spoken in relatively isolated Himalayan mountain pockets.

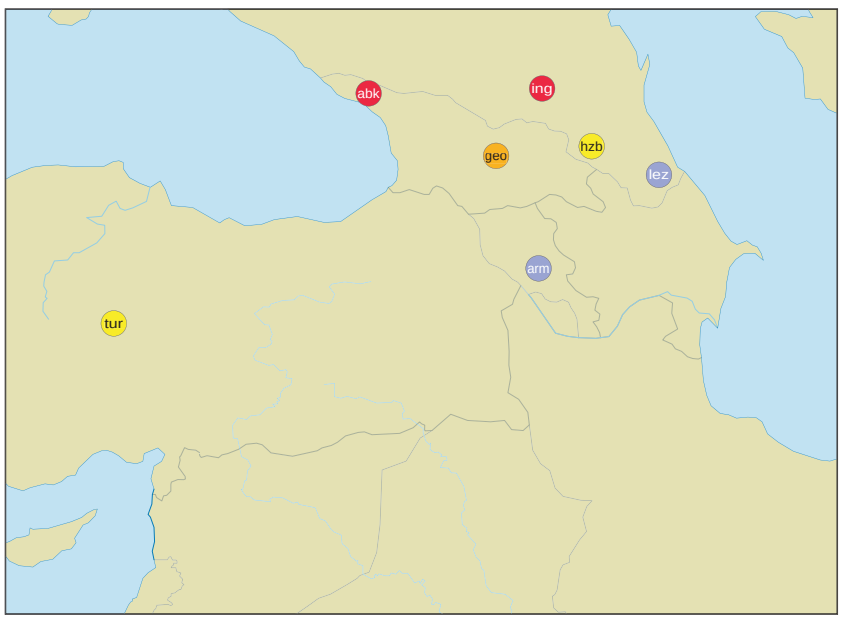
Acknowledgements

This research was supported by US National Science Foundation Grant No. 96-16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P. I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, Dave Peterson, and Suzanne Wilhite for help with data collection.

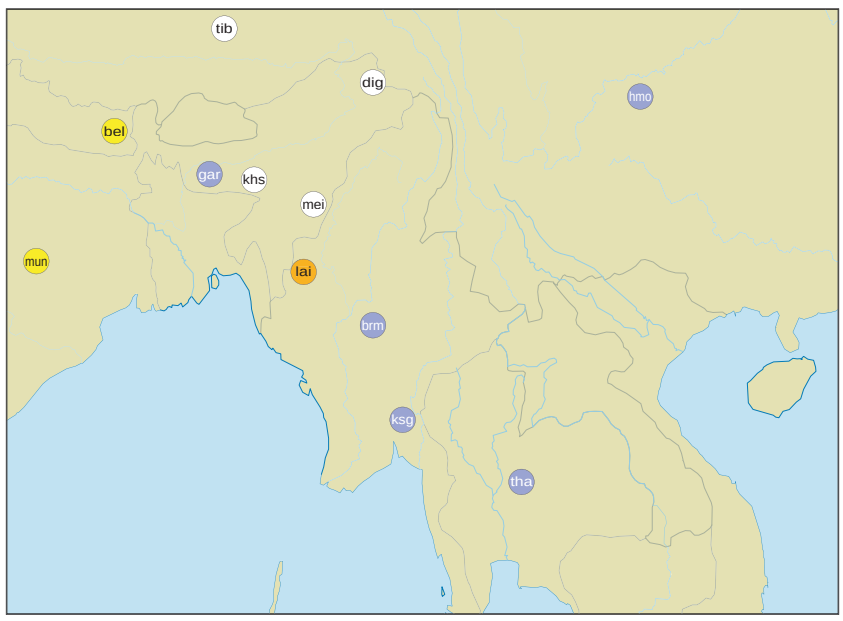
22 Inflectional Synthesis of the Verb

BALTHASAR BICKEL AND JOHANNA NICHOLS

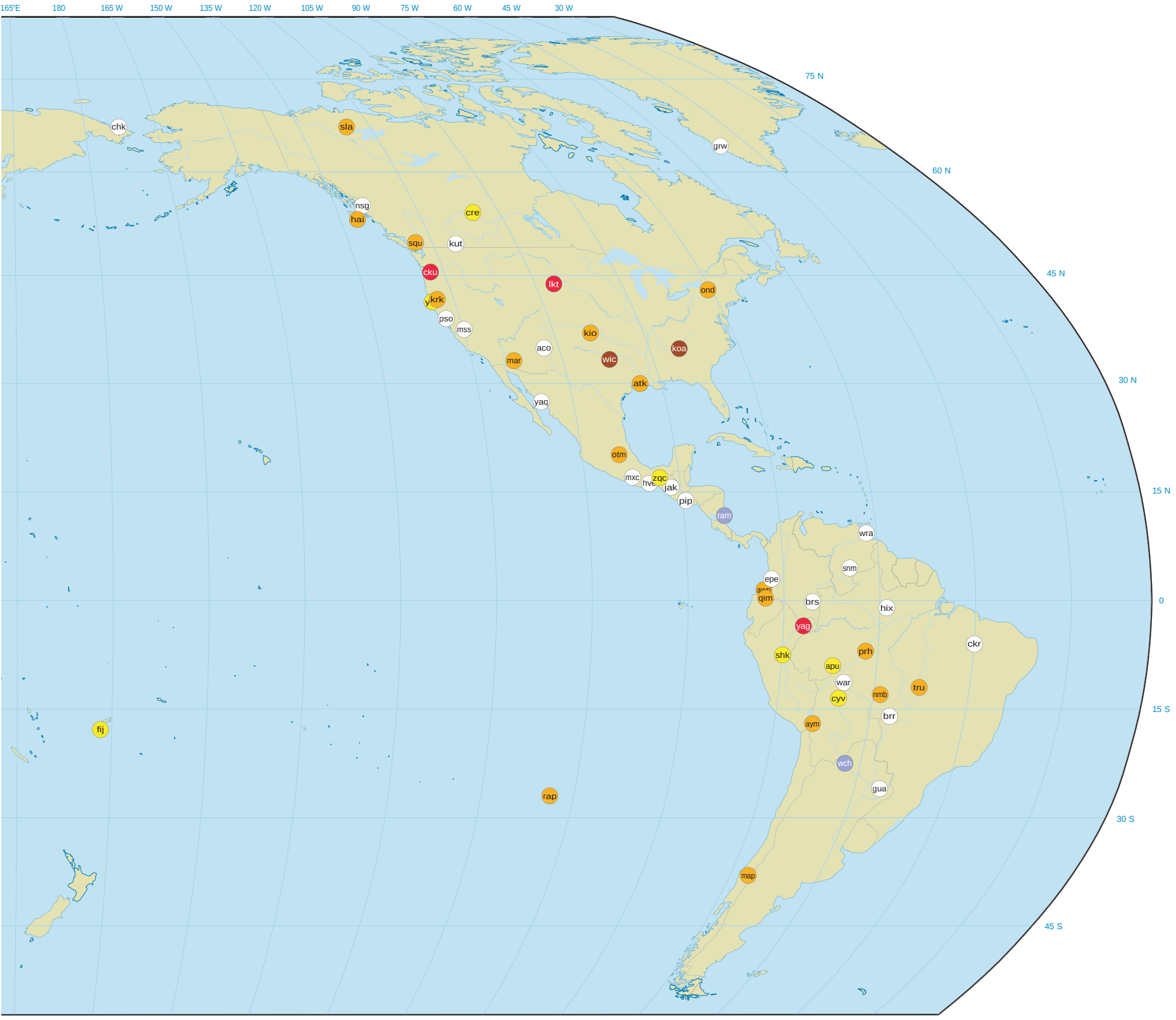


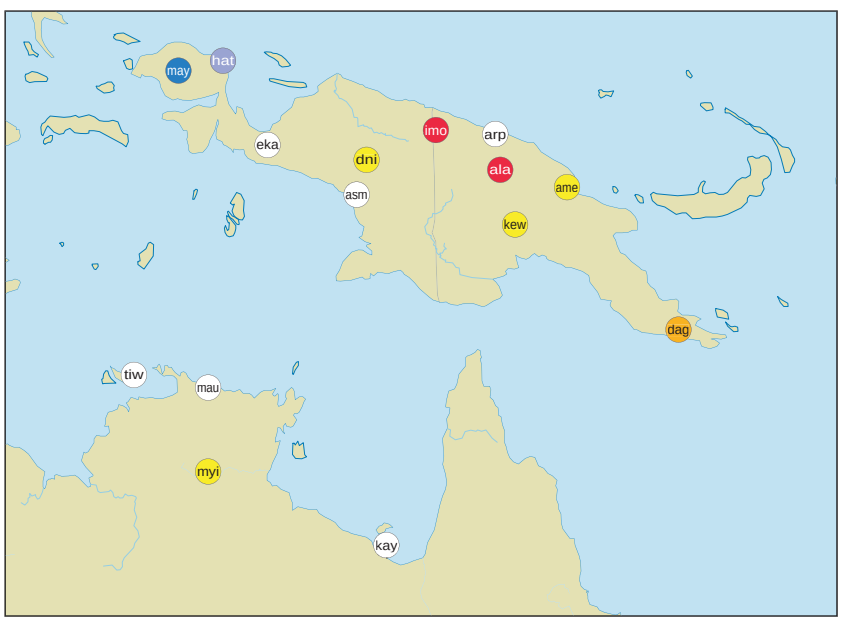


Caucasus and adjacent areas

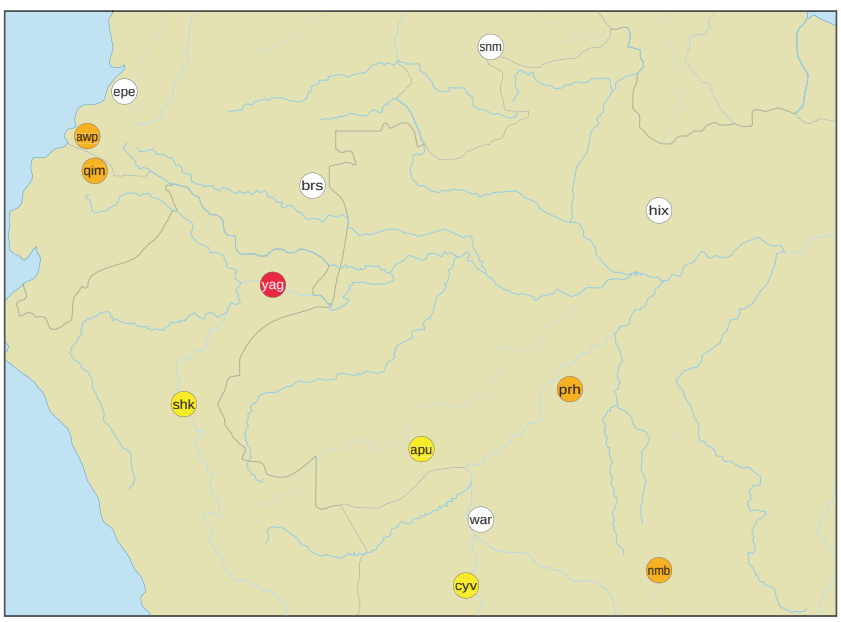


Myanmar and surrounding areas





New Guinea and northern Australia



northwestern South America

23 Locus of Marking in the Clause

JOHANNA NICHOLS AND BALTHASAR BICKEL

1 Definitions and values

Locus is a convenient one-word term for what is also known as *head/dependent marking*. In any kind of phrase, overt morpho-syntactic marking reflecting the syntactic relations within the phrase may be located on the head of the phrase, on a nonhead (i.e. on a dependent), on both, or on neither. In clauses, the arguments are dependents and the verb is the head. Examples of the locus types in clauses are (1)–(4). For ease of illustration these examples show the treatment of subjects and objects, in languages that treat both identically. As discussed below, not all languages treat subjects and objects identically, and where they are treated differently we have mapped the treatment of objects (specifically, direct objects (P below)). Not all languages treat pronoun and noun arguments identically, and where they are treated differently we have mapped the treatment of nouns.

The following locus types are distinguished on the map:

1. P is head-marked	71
2. P is dependent-marked	63
3. P is double-marked	57
4. P has no marking	42
5. Other types	2
total	235

1.1 Head marking. In (1), the verb agrees with the subject (3PL) and object (3SG), and their form and position show which is subject and which is object. The words 'boys' and 'rock' have no overt marking to indicate their subject and object functions. Thus all the marking of subject and object relations in this clause is on the verb, which is head of the clause.

(1) Tzutujil (Mayan; Guatemala; Dayley 1985: 282, 75)
jar aak'aalaa7 x-Ø-kee-k'aq aab'aj
the boys COMP-3SG-3PL-throw rock
pa rwi7 ja jaay
on top.of the house
'The boys threw rock(s) on top of the house.'

1.2 Dependent marking. In (2), the subject and object nouns 'old man' and 'firewood' bear cases marking their syntactic functions. (The absolutive case has a zero ending, but as this is the only zero ending in the case paradigm it is unambiguously identifiable as absolutive, hence it counts as a case.) The verb has no agreement with either subject or object. Thus all the marking of subject and object relations in this clause is on the subjects and objects themselves, which are dependents.

(2) Uradhi (Paman; Australia; Crowley 1983: 339)
wutpu-ŋku uma-Ø apa-n
old.man-ERG firewood-ABS pick.up-PST
'The old man picked up some firewood.'

1.3 Double marking. In (3), 'monkeys' and 'children' have case endings to indicate their syntactic functions, and the verb has affixes agreeing with both. Thus subject and object are marked twice, once on the subject and object themselves and once on the verb. This is double marking. (The third-person singular subject marker is zero, but is unambiguous as it contrasts with other person-number combinations in the paradigm.)

(3) Belhare (Tibeto-Burman; Nepal; Bickel, own fieldnotes)
kubaŋ-chi-ŋa pitcha-chi n-ten-he-chi
monkey-NSG-ERG child-NSG.ABS 3NSG.A-hit-PST-3NSG.P
'The monkeys hit the children.'

1.4 No marking. In (4), there is no marking of the syntactic relations on either the subject and object nouns or the verb. (The word order makes the syntax clear; but so it does in most of the above examples. Word order is not at issue in this chapter, which deals only with overt markers and their locations.)

(4) Thai (Nichols, own fieldnotes)
Daang hŭaróo Dam.
Dang laugh Dam
'Dang laughed at Dam.'

The examples of head, dependent, and double marking above have affixes or clitics as their marking. This is most common, but it is also possible for case and agreement markers to be isolating formatives, written and pronounced as separate words. The accusative marker of Dahalo (Cushitic; Kenya) is a separate word:

(5) Dahalo (Tosco 1991: 88)
ʔááta ʔéleto Dawa kabê
you know.NONPAST.2SG Dawa ACC
'Do you know Dawa?'

1.5 Other. The division into head, dependent, double, and zero marking does not exhaust the possible types. There are several low-frequency but systematic further patterns. One of them is **free (or floating) marking**, where the marker is positioned not on the head or the dependent of the phrase but on some other word in a position defined relative to the head or to the phrase boundaries. For example, in Yagua (Peba-Yaguan; Peru), an overt object NP, if roughly definite (see Payne 1990: 255 and 364–7 for more on the pragmatic conditions), is marked by a clitic which attaches to the preceding word. This is floating because it is controlled by one word (the object) but located on another (the preceding word, whatever it may be). The following examples show the object preceded by the verb, the subject, and an adverbial word, respectively.

(7) Yagua (Payne 1990: 255) (I, II are agreement classes)
a. *siimyi=ñíí quiiv-ą*
3SG.I.eat-3SG.II fish
'He is eating the fish.'
b. *siimyiy Alchíco=níí quiiv-ą*
3SG.I.eat Alchico-3SG.II fish
'Alchico is eating the fish.'
c. *siimyiy sinumu=níí quiiv-ą*
3SG.I.eat land.LOC-3SG.II fish
'He is eating the fish on land.'

Rather than typologize clause locus based on the treatment of all three of subjects, direct or primary objects, and indirect or secondary objects, we have relied on just the treatment of direct or primary objects in this map. The reasons for this choice, among various alternatives, are set forth in §3 ("Defining locus types") below. In cases of differential object marking (Bossong 1985, Aissen 2003), where some objects are zero-marked and some (generally definite or nongeneric objects) are overtly marked, it is not always straightforward to determine which is the default form. Here we selected the overt marking type (generally an accusative or dative case) as the survey object.

2 Geographical distribution

The head-marked clause is common in the Americas and Australia-New Guinea and very rare elsewhere. (There are five tokens in Africa, but three of them come from the young but widespread Bantu family.) The dependent-marked clause is common in Eurasia and northern Africa, sparsely but solidly attested in South America,

and rare in North America. In Australia-New Guinea it occurs in two clusters: the eastern highlands of New Guinea and the south, east, and interior of Australia. The Australian cluster is genealogically biased, with most of the tokens coming from the very old Pama-Nyungan family; the one in New Guinea is genealogically better distributed but smaller. Double marking is moderately well attested in the Americas, Australia-New Guinea, and the southern fringe of Eurasia (chiefly in the Caucasian and Himalayan mountain enclaves), and seems to be favoured particularly in Australia and the westernmost Americas. The zero-marked object is, unsurprisingly, common in South-East Asia and western Africa, two well-known centres of morphological simplicity; but it is also very common in New Guinea and moderately common in eastern Africa and Central to South America, among languages of average or higher morphological complexity.

3 Defining locus types

Locus types, whether for particular constituents or as whole-language types, have generally been defined by considering the locus marking of several different relations. Nichols 1986 defined whole-language types by considering marking of possessive noun phrases with noun possessors (e.g. *neighbour's house*) and with pronoun possessors (e.g. *her house*); subject, direct object, and indirect object relations in the clause, all three with both noun and pronoun arguments; adpositional phrases with both noun and pronoun objects; and noun phrases with attributive adjectives. Nichols 1992 removed adpositional phrases because of surveying difficulties but kept the others, which added up to nine different syntactic locations per language at which there might be head marking, dependent marking, both, or neither. Whole-language types were then defined by various elementary mathematical operations on these nine: the sum of head-marking points minus the sum of dependent-marking points, or the ratio of head- to dependent-marking points. The head-marking type was the lowest one-third of the head/dependent-marking ratios, or the lowest sums. Cysouw 2002 considers head marking and dependent marking as separate parameters, and uses the sum for each to indicate types.

These numerical operations have various disadvantages (see Cysouw 2002). More importantly, the range of nine morphosyntactic points on which they are measured presents some difficulties. All nine are not equally likely to have one or another marking type, as basing types on their sum or ratio would imply. Rather, some syntactic relations are especially prone to one or another kind of marking. For instance, the marking of the S or A argument is prone to reflect universals more than type: agreement with this argument is common even in otherwise dedicated dependent-marking languages, and its inflection is likely to be driven by topicality-related matters at least as much as by morphosyntactic type. Indirect or secondary objects are generally not core arguments and are very often treated as obliques, which are almost always dependent-marked. Pronominal arguments are likely to either inflect for case in otherwise caseless languages or cliticize to their heads (or both) regardless of the overall clause locus type of the language. Attributive adjectives are very rarely head-marked or double-marked. Furthermore, parts of speech behave differently: pronouns are likely to cliticize to verbs when nouns do not, and pronouns are likely to make case distinctions when nouns do not. These are universal tendencies.

Though the typologist obviously needs all of these kinds of information in order to determine preferences for different kinds of marking in different parts of the grammar, for purposes of determining the head/dependent-marking type of a language the nine-point survey muddies the picture, as marking at several of the points reflects universal tendencies rather than the marking type of the individual language. By removing the points most prone to follow universal tendencies, we can distill the morphosyntactic range down to two phrase types: possessive phrase with noun possessor (e.g. *neighbour's house, the colour of grass*) and direct or primary object in the transitive clause (e.g. *wrote books, broke a glass*). The whole-language type can be defined on the noun phrase and the clause together. This exemplar-based survey of marking types makes it possible to map out the four cardinal types (head marking, dependent marking, double marking, zero marking) clearly and simply.

We created a preliminary set of maps using various definitions of the marking types (nine-point survey, smaller surveys, and the exemplar-based definition used here) and found that in all of them the same (or much the same) geographical areas showed up as favouring various types. Thus the pared-down, exemplar-based approach loses no information and in fact appears to gain some information in that (especially for the whole-language types) areas stand out more sharply.

Even when considering just a single exemplar, questions can arise as to how to classify nonprototypical types. If a language had two different kinds of marking for one morphosyntactic point (e.g. most objects trigger verb agreement but some do not, animate and inanimate nouns have different case inventories or different privileges of occurrence as subject of transitive verb), we chose the majority or open or default pattern; if there was no clear majority/default/open pattern we coded the marking as split. If first- and/or second-person agreement markers differ from third-person ones, we have included only the third-person form; and likewise for any enumerable or delimitable special forms of marking.

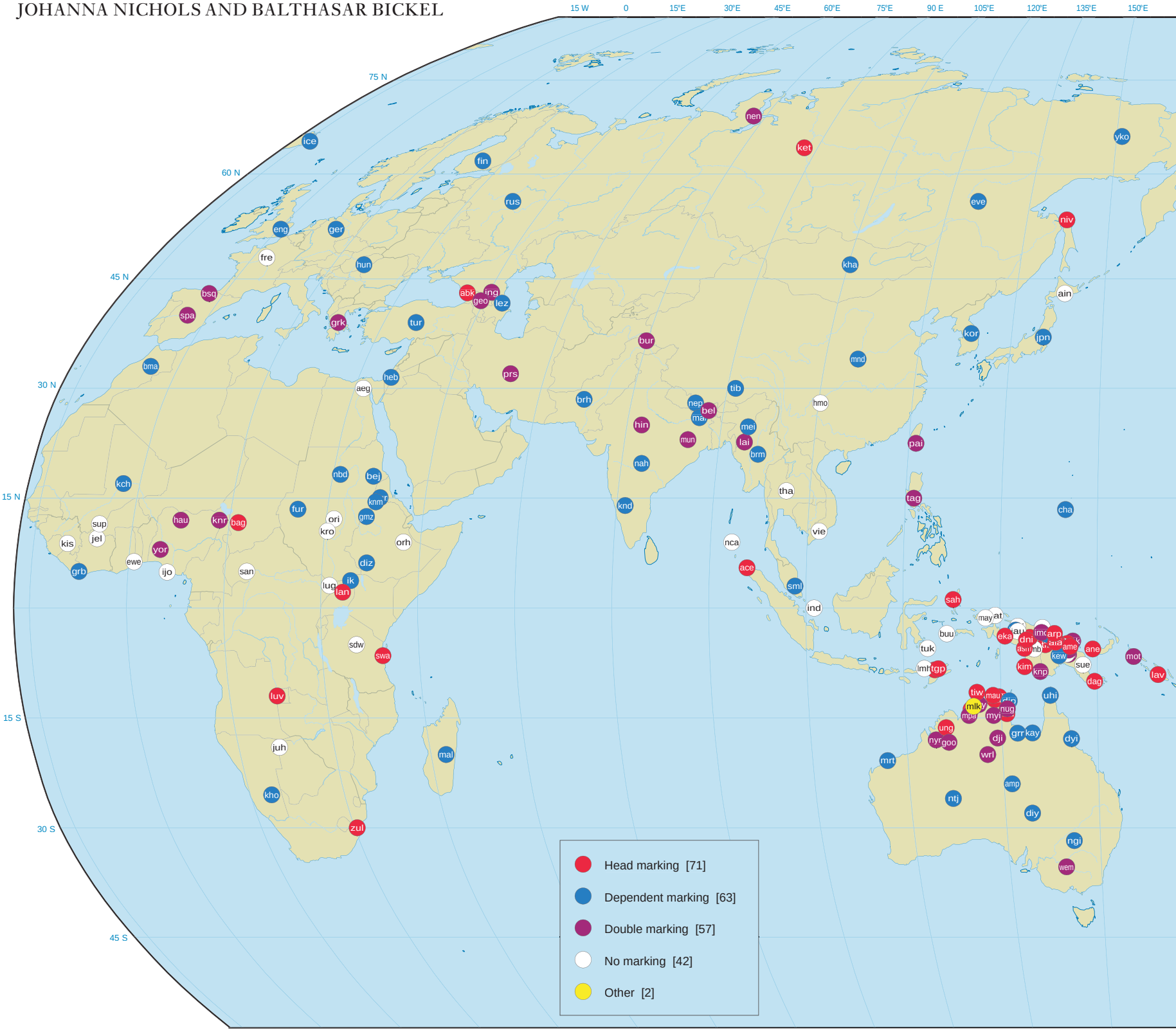
Other Maps using exemplar-based definitions include 20–22, 24, and 25. For more discussion of the exemplar-based method see Bickel and Nichols 2002.

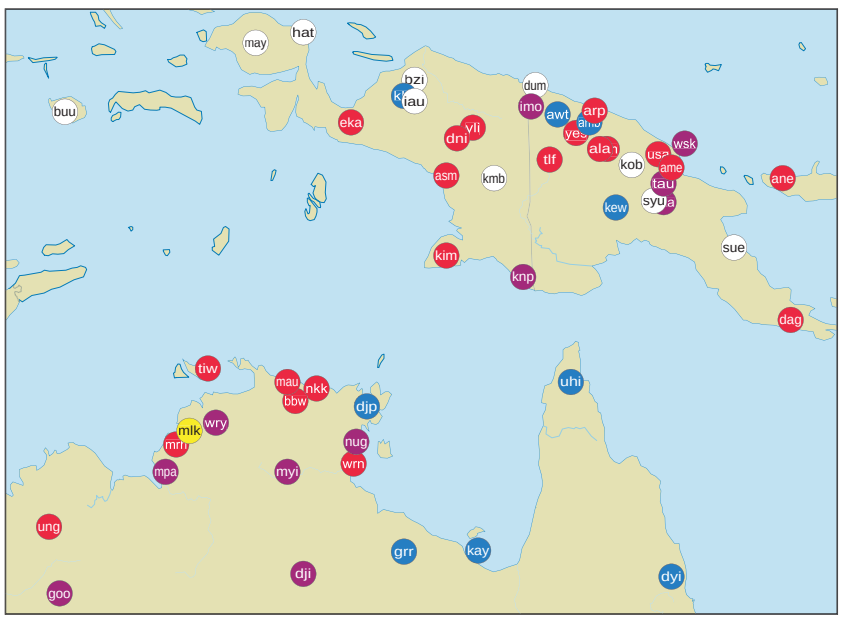
Acknowledgements

This research was supported by US National Science Foundation Grant No. 96-16448 (Nichols, P .I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P. I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, David A. Peterson, Rebecca Voll, and Suzanne Wilhite for help with data collection.

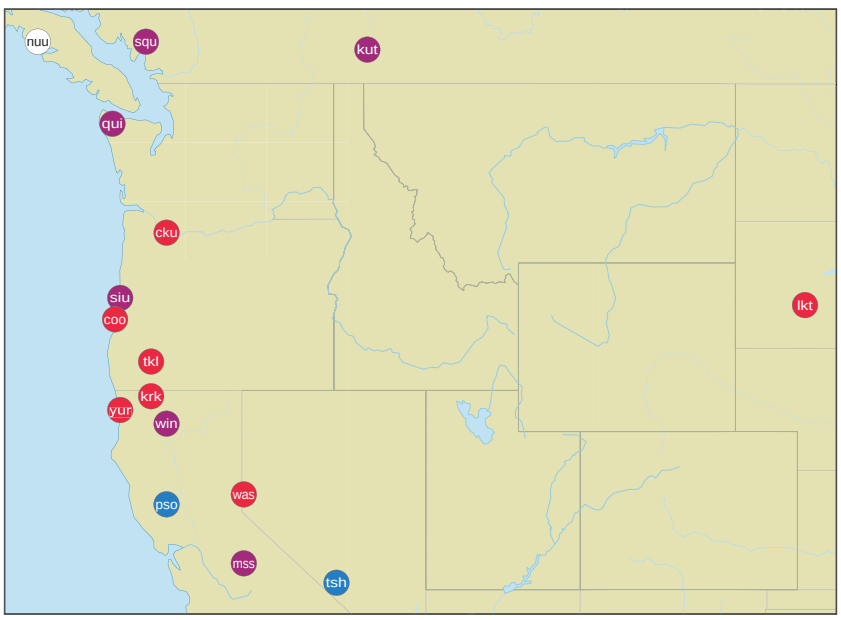
23 Locus of Marking in the Clause

JOHANNA NICHOLS AND BALTHASAR BICKEL

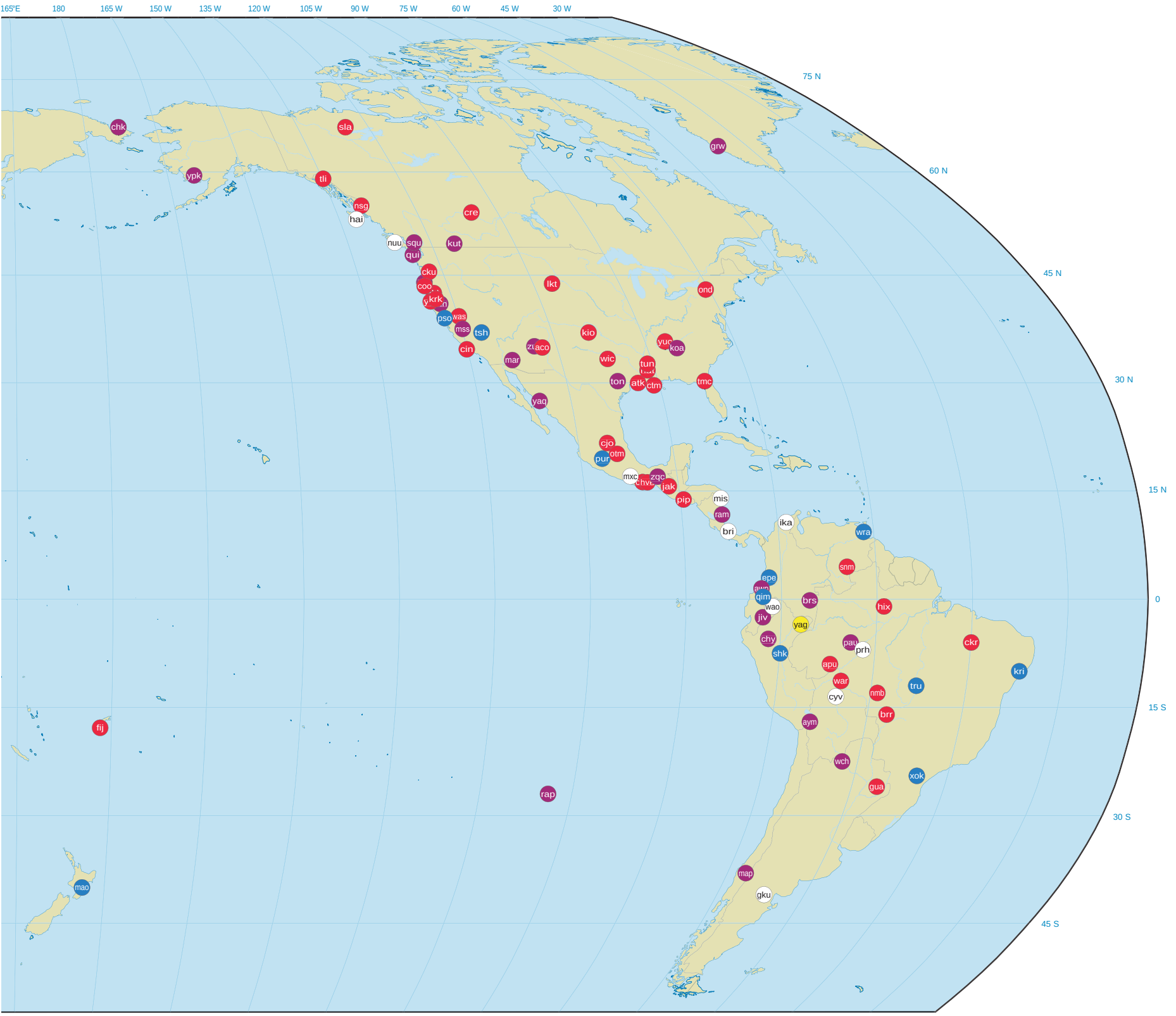




New Guinea and northern Australia



western United States



southern Mexico and Central America

northwestern South America

24 Locus of Marking in Possessive Noun Phrases

JOHANNA NICHOLS AND BALTHASAR BICKEL

1 Definition and values

Locus is a convenient one-word term for what is also known as *head/dependent marking*. In any kind of phrase, overt morphosyntactic marking reflecting the syntactic relations within the phrase may be located on the head of the phrase, on a nonhead (i.e. on a dependent), on both, or on neither. Examples illustrating marking of various kinds in possessive noun phrases are (1)–(4) below. In possessive phrases, the possessed noun is head and the possessor is dependent.

The following locus types are distinguished on the map:

1.	Possessor is head marked	77
2.	Possessor is dependent marked	98
3.	Possessor is double marked	22
4.	Possessor has no marking	32
5.	Other types	6
	total	235

These five types are illustrated in the following subsections.

1.1 Head marking.

(1) Acoma (Keresan; New Mexico; Miller 1965: 177)
s'adyúm'ə *gâam'a*
1SG.brother 3SG.house
'my brother's house' (lit. 'my-brother his-house')

In these examples the possessed noun (the head) agrees in person and number with the possessor noun, the most common pattern for head-marked noun phrases. Agreement in gender is also fairly common. A few languages have a non-agreeing marker on the head noun. In Fijian, a possessive affix *-i* marks possessed nouns; that it does not vary for person or number of the possessor is shown in (2b):

(2) Fijian (Austronesian; Fiji; Dixon 1988: 36)
a. *a* *mata-i* *Jone*
ART eye-POSS John
'John's eye'
b. *a* *liga-i* *'eirau*
ART hand-POSS 1DU.EXCL
'our hand(s)'

In Yoruba, the possessed noun has its final vowel lengthened on a mid tone before a consonant-initial possessor noun:

(3) Yoruba (Benue-Congo; Nigeria; Awobuluyi 1978: 40)
owóo *Dàda*
money.POSS Dada
'Dada's money' (cf. *owó* 'money')

1.2 Dependent marking. In the following examples from Chechen (Nakh-Daghestanian), the possessor noun is in the genitive case.

(4) Chechen (Nichols, own data)
a. *loem-an* *k'orni*
lion-GEN baby.animal
'lion cub', 'lion's cub' (lit. 'of-lion cub')
b. *mashien-an* *maax*
car-GEN price
'the price of a car' (lit. 'of-car price')

1.3 Double marking.

(5) Southern Sierra Miwok (Miwok-Costanoan; California; Broadbent 1964: 133)
cuku-ŋ *hu:ki-ʔ-hy:*
dog-GEN tail-3SG
'dog's tail' (lit. 'of-dog its-tail')

1.4 No marking.

(6) Tiwi (isolate; northern Australia; Osborne 1974: 74)
jərəkəpai *tuwaɹa*
crocodile tail
'(a/the) crocodile's tail' (lit. 'crocodile tail')

(7) Asmat (Asmat-Kamoro; Papua, Indonesia; Voorhoeve 1965b: 136, 133)
a. *Warsé* *ci*
Warse canoe
'Warse's canoe' (lit. 'Warse canoe')
b. *no* *cém*
1SG house
'my house' (lit. 'I house' or 'me house')

In the Acoma examples in (1), the possessed noun (the head) is inflected, agreeing in person and number with the possessor. In the Chechen examples, in contrast, the possessor noun (the nonhead) is inflected, bearing the genitive case. In Southern Sierra Miwok both words are inflected. In the Tiwi and Asmat examples there is no inflection of either word; possessive phrases are formed by juxtaposition of the uninflected possessor and possessed nominals. (This is not the only adnominal construction of Tiwi; see the next paragraph.)

In these examples the inflection is affixal. However, the same work can also be done by separate words. Head marking using a separate word is illustrated by (8) from Tiwi, where the marker of possession is the uninflected pronoun *ŋara* 'he'. That it is syntactically attached to the head noun 'tail' and not the possessor noun 'crocodile' is shown when the order of possessor and possessed nouns is inverted: in (6a–b) 'he' immediately precedes 'tail' regardless of the latter's position relative to 'crocodile'.

(8) Tiwi (Osborne 1974: 74–5)
a. *jərəkəpai* *ŋara* *tuwaɹa*
crocodile he tail
'the crocodile's tail'
b. *ŋara* *tuwaɹa* *jərəkəpai*
he tail crocodile
'id.'

Dependent marking using a separate word is illustrated by (9), where *na* is a postposition, and by English phrases with *of*, e.g. *the price of oil*.

(9) Amele (Trans-New Guinea; Papua New Guinea; Roberts 1987: 139)
Naus *na* *jo*
Naus of house
'Naus's house'

1.5 Other. The division into head, dependent, double, and zero marking does not exhaust the possible types. There are several low-frequency but systematic further patterns, all of them grouped together as "Other" on the map. One of them is **free** (or **floating**) **marking**, where the marker is positioned with respect not to the head or the dependent of the phrase but relative to the phrase boundaries. The commonest such position is probably second, or Wackernagel, position, with the marker following the first word or similar unit of the phrase, as in the following NP examples from Chamorro, where *=n* is a second-position clitic following the first tonic word in the phrase ("=" is a clitic boundary).

(10) Chamorro (Austronesian; Guam; Topping 1980: 223)
i=lepblo=n *estudiante*
ART=book=LINK student
'the student's book'

That the *=n* is in second position is shown by the fact that in phrases with attributive modifiers, where word-order inversion is possible, it follows whichever word comes first:

(11) Chamorro (Topping 1980: 208)

a. *i=dánkolo=n* *taotao*
ART=big=LINK man
'the big man'

b. *i=kareta=n* *Japanese*
ART=car=LINK Japanese
'Japanese car'

Another minor type is **headward-migrated dependent marking**, where a fully inflected dependent (typically a pronominal argument) cliticizes to the head, as in Bororo (Macro-Ge; Mato Grosso, Brazil):

(12) Bororo (Crowell 1979: 197)

barae *eno* *moto*
Brazilians 3PL.GEN land
'Brazil' (lit. 'Brazilians' land')

This is different from (8) in that the pronominal piece in (12) is case-inflected and is therefore a syntactic word, while that in (8) has no case and can therefore be regarded as a phonologically word-like grammatical formative.

2 Exemplar-based survey

Though a number of languages have more than one adnominal construction, for this map we sought just one type per language. We have mapped only possessive constructions with overt heads and noun possessors. For all languages in which there was more than one marking pattern for possessive NPs, we chose whichever is default or has the fewest restrictions. For instance, if a closed or restricted lexical group of nouns (such as kin terms and/or body parts or "inalienables": see Chapter 59 on possessive classification) has a distinct form of possessive marking, we have not included this but have mapped only the treatment of the open, or default (or "alienable"), set of nouns. If nouns and pronouns have different marking we have mapped only the noun marking; if only first and second persons have a distinctive form of marking we have not included it; and likewise for any enumerable or delimitable special forms of marking. Though kin terms and body parts are prototypically possessed and might be thought the best possible examples of possessive phrases, in the languages that formally differentiate "alienable" from "inalienable" possession it is always the "inalienables" that are the closed or defined class. Hence they are not the open, default construction we seek in this survey. Furthermore, the open class of possession reflects the language's typological preferences for marking locus, whereas "inalienable" possession shows a clear worldwide preference for head marking. More generally, we found that closed-class and minority patterns are so driven by universal preferences in their marking as to yield little interesting typological or geographical variation: pronouns often inflect for case or are otherwise dependent-marked in languages that have no cases on nouns (see Chapter 50), and they are also prone to cliticize to heads when nouns are not.

The result of surveying only open-class marking is fewer data points than have been used to define locus types in earlier literature (e.g. Nichols 1986, 1992; Cysouw 2002). This is an *exemplar-based* survey which takes the most typologically revealing and least universal-driven pattern to represent the language's type. In addition to greater typological clarity it also produces a clearer picture of continental and larger geographical distributions, albeit at the price of a more schematic description of each individual language's morphosyntax. (For further discussion of the examplar-based method, see §3 "Defining locus types" in Chapter 23, and Bickel and Nichols 2002; for other applications, see Chapters 20–22, and 25.)

3 Geographical distribution

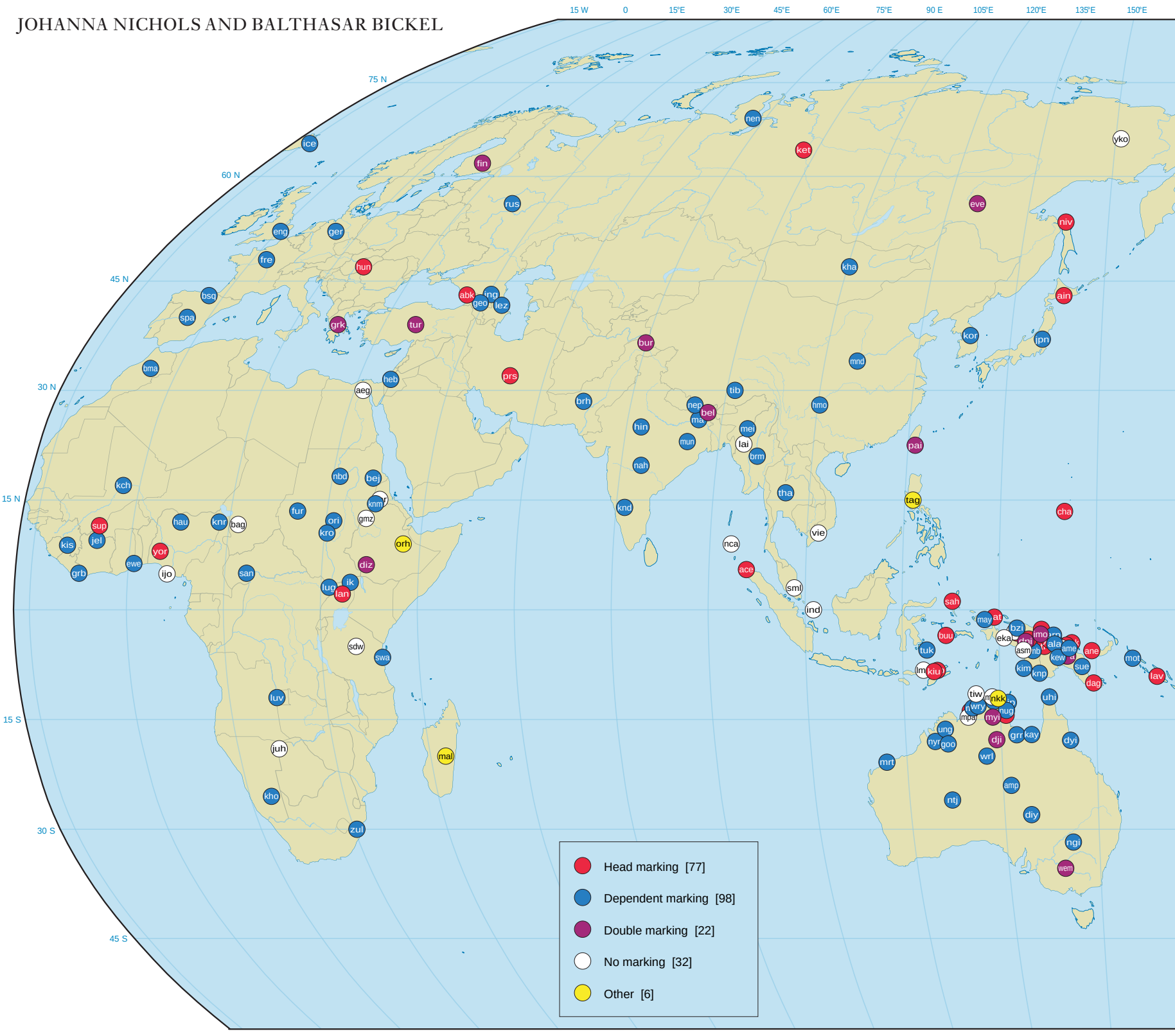
Head-marked possessive NPs are common in the Americas and the Pacific (chiefly Melanesia) and infrequent elsewhere. Dependent-marked NPs have a roughly complementary distribution to this: they are frequent in all parts of Africa, Eurasia, and Australia-New Guinea. The only area where the two types overlap appreciably is in New Guinea. Double-marked possession is rare; it is found around the Eurasian periphery, in the Himalayas, and along the Pacific coast of North America (see Chapter 22 on synthesis for a similar distribution). Zero-marked possession is also uncommon; instances of it on the map are mostly found near the equator, but apart from this curious distribution it does not form any true clusters.

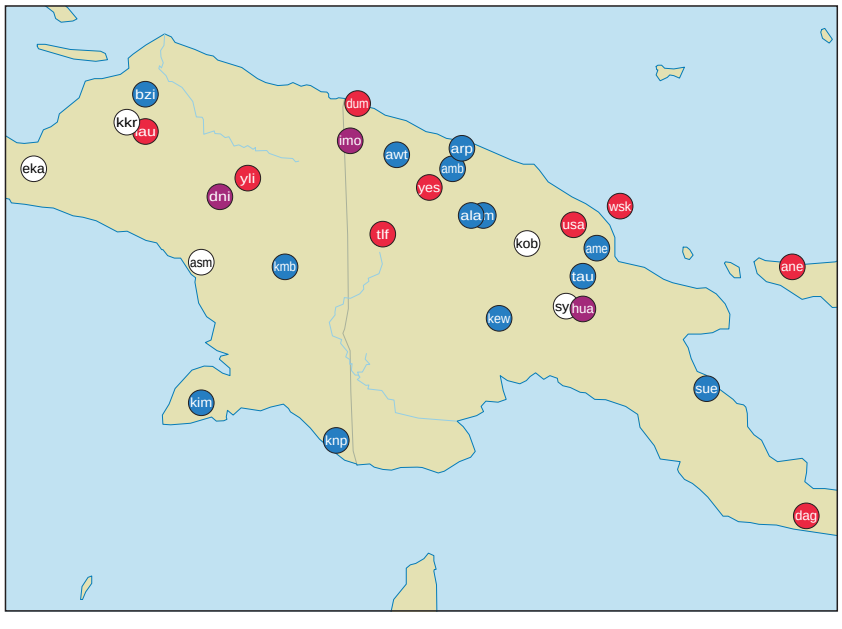
Acknowledgements

This research was supported by US National Science Foundation Grant No. 96-16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P.I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, David A. Peterson, Rebecca Voll, and Suzanne Wilhite for help with data collection.

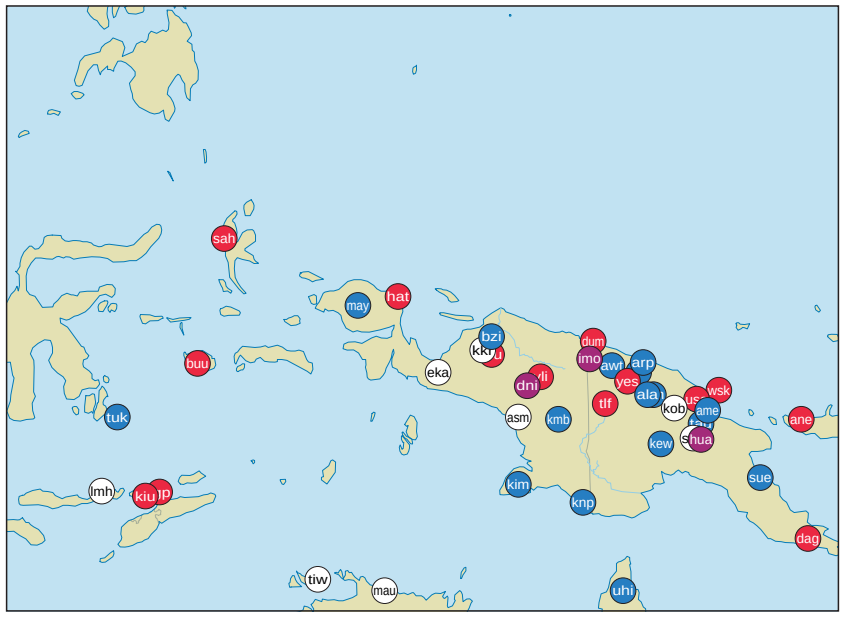
24 Locus of Marking in Possessive Noun Phrases

JOHANNA NICHOLS AND BALTHASAR BICKEL

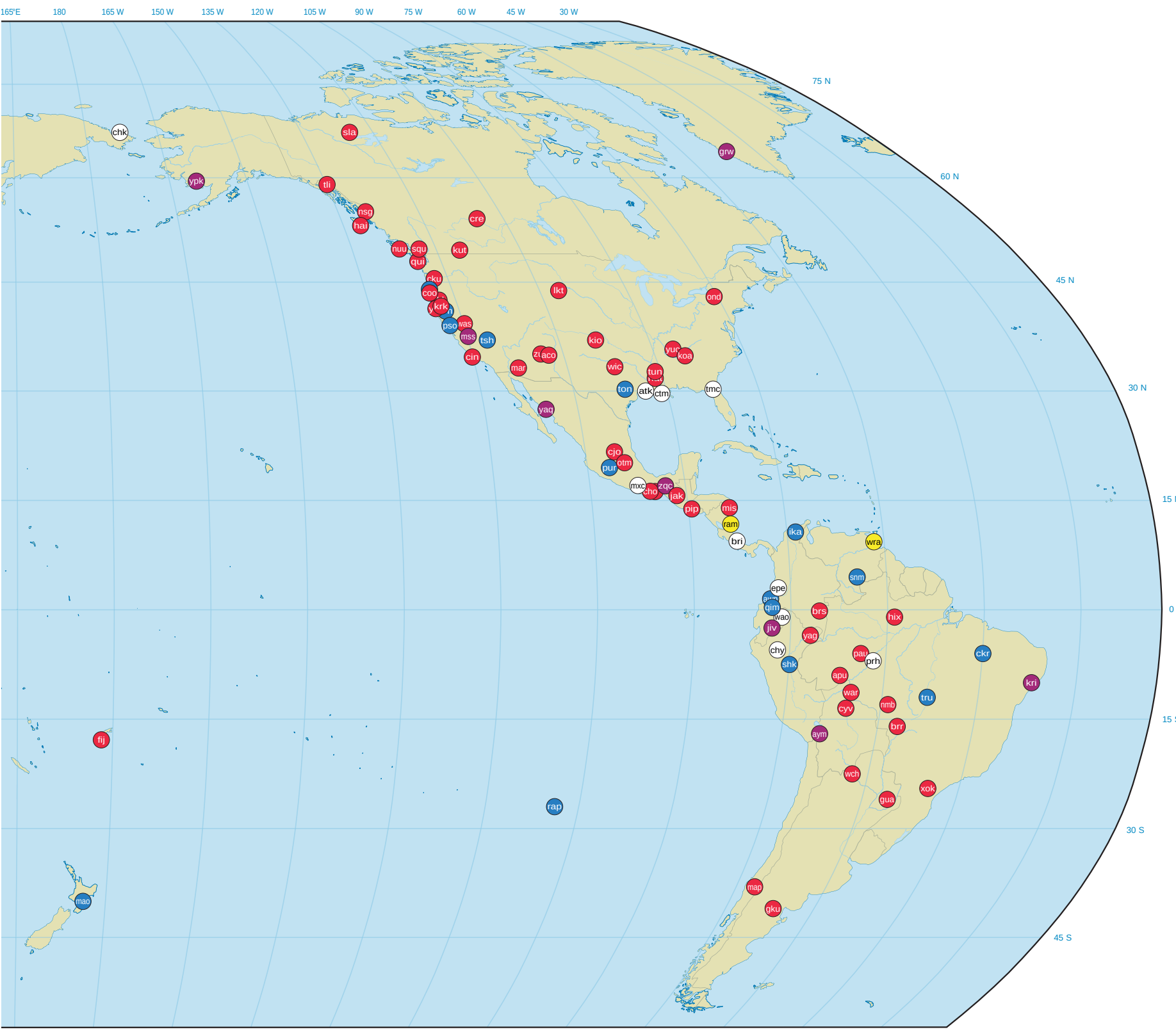


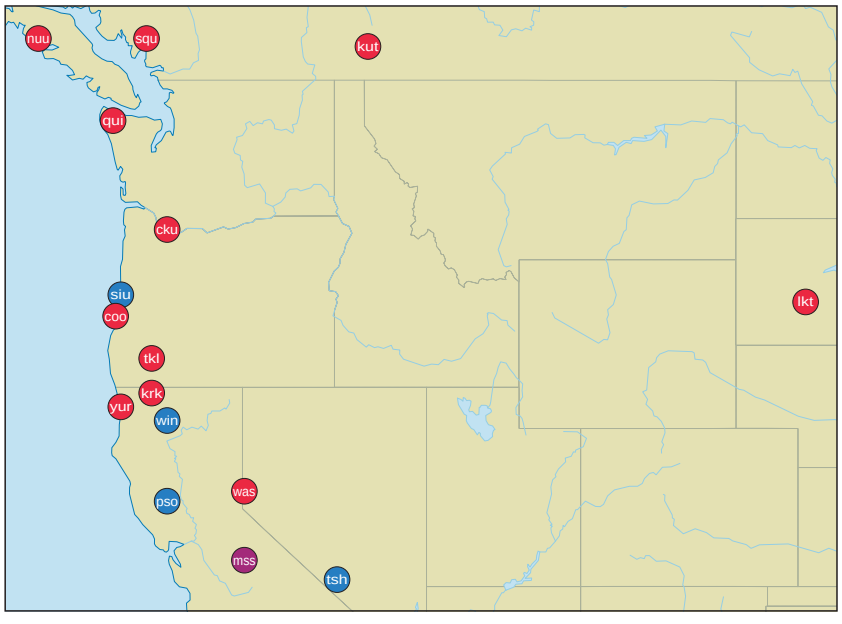


New Guinea

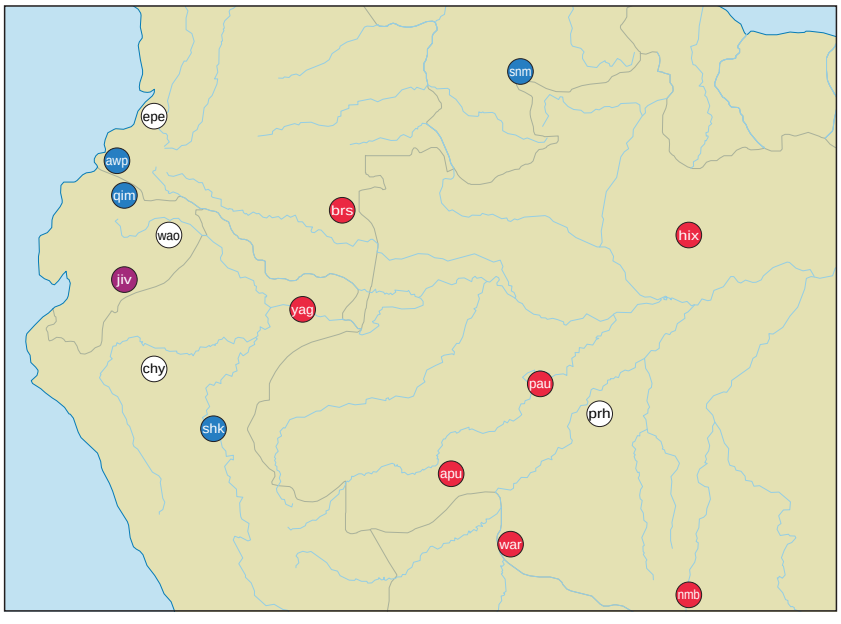


northern Australia





western North America



northwestern South America

25 Locus of Marking: Whole-language Typology

JOHANNA NICHOLS AND BALTHASAR BICKEL

1 Definitions and values

Locus is a convenient one-word term for what is also known as *head/dependent marking*. In any kind of phrase, overt morpho-syntactic marking reflecting the syntactic relations within the phrase may be located on the head of the phrase, on a nonhead (i.e. on a dependent), on both, or on neither. In possessive phrases, the possessed noun is head and the possessor is dependent. Examples illustrating marking of various kinds in possessive phrases are given in Chapter 24. In clauses, the arguments are dependents and the verb is the head. Examples of the locus types in clauses are given in Chapter 23. The division into head, dependent, double, and zero marking does not exhaust the possible types; see the same chapters for examples of additional minor locus patterns.

The maps in this chapter are based on the locus of marking in both the possessive NP and the clause. More specifically, we looked at marking of:

- Possessor and possessed in the NP. If there was more than one marking pattern, we chose whichever is default or has the fewest restrictions. (See Chapter 24.)
- Direct or primary object (i.e. P) in the transitive clause. Again, it is only the default form of marking that is mapped. (See Chapter 23.)

Each of these two studies is an *exemplar-based* survey which seeks a cardinal representative of locus marking for each relation, a representative which maximally reflects type and minimally reflects universals. (For discussion of the exemplar-based approach see §3 "Defining locus types" in Chapter 23.) We combine these two here to produce an exemplar-based definition of whole-language types, which is simply the intersection of the definitions for NP and clause locus: the consistently head-marking type has head-marked P and possessor, and analogously for the consistently dependent-marking and double-marking types. That is, a language displaying the cardinal type shows the same kind of marking in both relations.

1. Consistently head marking	47
2. Consistently dependent marking	46
3. Consistently double marking	16
4. Consistently zero marking	6
5. Inconsistent marking or other type	120
total	235

Examples of the various types follow.

*1.1 **Head-marking language***, with head marking in both NP and clause:

(1) Tzutujil (Mayan; Guatemala; Dayley 1985: 286)

a. *r-k'aajool* *n-b'esiino*
3SG-son 1SG-neighbour
'my neighbour's son'

b. *jar* *aak'aalaa7* *x-Ø-kee-k'aq* *aab'aj*
the boys COMP-3SG-3PL-throw rock
pa *rwi7* *ja* *jaay*
on top.of the house
'The boys threw rock(s) on top of the house.'

*1.2 **Dependent-marking language***, with dependent marking in both NP and clause:

(2) Chechen (Nakh-Daghestanian; Nichols, own data)

a. *loem-an* *k'orni*
lion-GEN baby.animal
'lion cub', 'lion's cub' (lit. 'of-lion cub')

b. *mashien-an* *maax*
car-GEN price
'the price of a car' (lit. 'of-car price')

c. *oox-a* *cha-ca* *txou-Ø* *qeelira*
1PL-ERG straw-INSTR roof-ABS cover.PST
'We thatched the roof/we covered the roof with straw.'

*1.3 **Double-marking language***, with case marking on possessors and objects as well as agreement on possessed nouns and verbs. In the following examples the relevant markers are in boldface.

(3) Mangarrayi (northern Australia; Merlan 1982: 30, 64)

a. *Ø-ṇi-ṇawu* *ṇa-muṛimuṛi-wuṇyaŋ-**gu**-bayi*
N.ABS-name-his M.NOM-father's.father-3PL-GEN-FOC
'their paternal grandfather's name'

b. ***ṇan**-yaba* *ga-**ŋa**-ḍaṇidba*
M.ACC-brother TNS-1SG>3SG-wait
'I'm waiting for my brother.'

(4) Jivaro (Jivaroan; Peru; Pellizzaro 1969: 14–15, Gnerre 1999: 61)

a. *wi-**na*** *nua-**r***
1SG-GEN woman-1SG
'my wife'

b. *tarách-ru-mi-**n*** *nawánt-**an*** *su-**Ø**-sá-ja-i*
garment-POSS-2SG-OBJ son-OBJ give-3-ASP-1SG-DECL
'I gave your garment to (my) son.'

In (4b), the zero suffix is unambiguously third person and fits into a larger paradigm whose other members are overt, so it counts as a kind of marking.

1.4 Zero marking. In contrast, in a zero-marking language there is no marking on either object or possessor, for third person singular or any other person-number combination:

(5) Pirahã (Mura; Brazil; Everett 1986: 272, 201, 206)

a. *ti* *kaíi*
1 house
'my house'

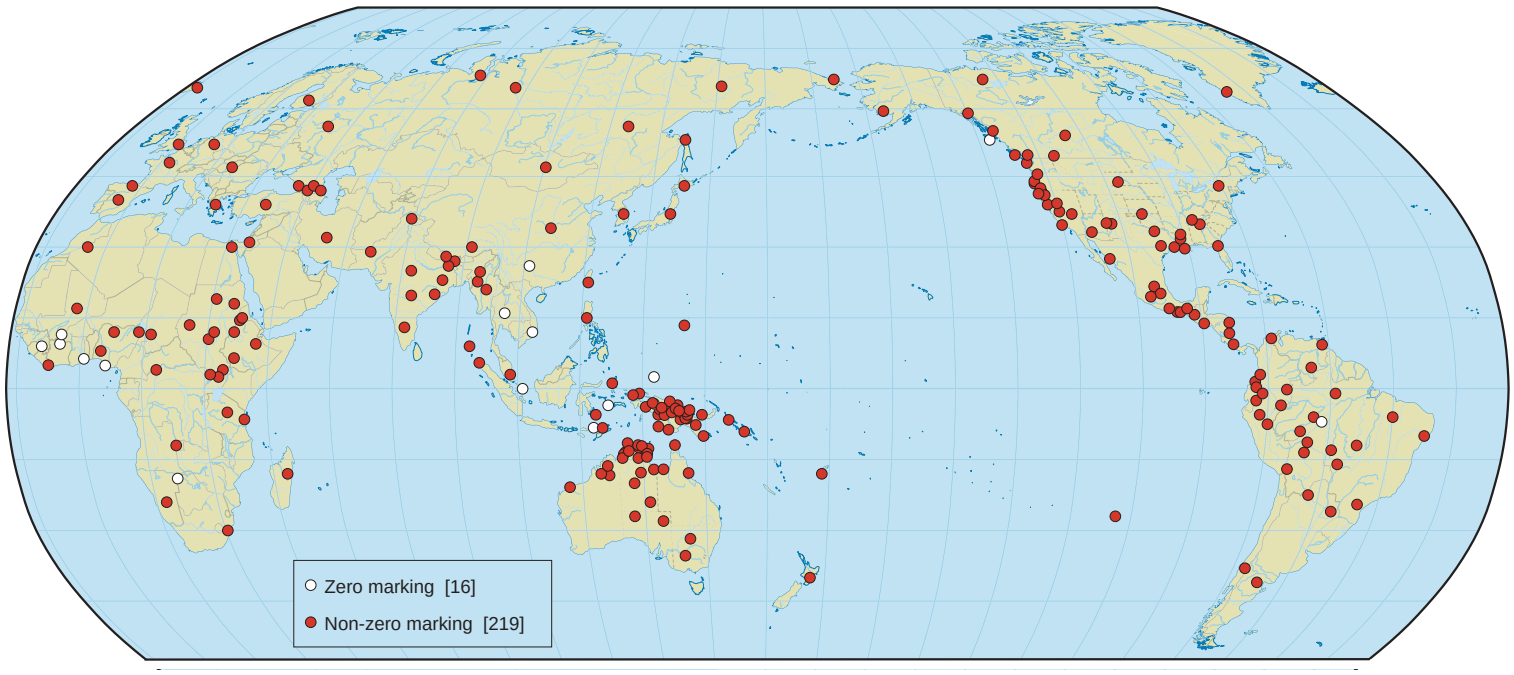
b. *ti* *xíbogi* *ti-baí*
1 milk drink-INTENSIFIER
'I really drink milk.'

c. *ti* *gí* *kapiigaxiítoii* *hoa-í*
1 2 pencil give-PROX
'I give the pencil to you.'

2 Geographical distribution

The exemplar-based definitions used here pick up only the most committed languages, as it were, of each type, giving geographical distributions that are the set intersection of those generated from just one syntactic relation and hence sparser but cleaner and more useful to historical linguistic geography. The distributions of the head-marking and dependent-marking types are sharper versions of those shown in Chapters 23 and 24: the head-marking type is well attested in the Americas and Melanesia but vanishingly rare elsewhere; the dependent-marking type is strongest in Africa, Eurasia, and perhaps Australia-New Guinea and infrequent (though not rare) elsewhere. The double-marking type is rare overall; its decimated distribution, the set intersection of double marking in Chapters 23 and 24, is a Pacific Rim distribution showing an enclave effect in the Himalayas, and one outlier in Modern Greek.

Only the zero-marking type has a distribution that is not simply a thinner version of that in Chapters 23 and 24: zero-marked P and



Map 25A Zero Marking of A and P Arguments

zero-marked possessors appear in languages of very different types and usually not in the same languages, perhaps because zero-marked possession is much like compounding or incorporation, phenomena often associated with morphological complexity, while zero-marked clause arguments are part of the signature of the isolating morphological type. Only six languages fall into the zero-marking type as defined on P and possessor. Hence zero marking as shown on the map here has a scant and vaguely southern distribution of no particular areality.

The inset map shows zero marking defined entirely on the clause, as zero marking of both the A argument and the P argument.

1.	Zero marking	16
2.	Other (non-zero marking)	219
	total	235

Values of Map 25A Zero Marking of A and P Arguments

Here the geographical result is two prominent clusters in well-known hotbeds of the isolating morphological type, West Africa and South-East Asia, and a distribution that is overall a sparser version of those in Chapters 23 and 24. (Defining the whole-language type on A and P is not typologically very satisfactory, as it properly describes only the clause. For all four cardinal types on the main map, and not just for the zero-marking type shown in the inset, the number of languages falling into the type as defined for A and P is larger than as defined for P and possessor.)

About half of the languages in our sample fall into one of these four consistent locus types; the others have different kinds of marking for each of the two syntactic relations that define the types. Of the four whole-language types, the polar ones—the consistent head-marking and consistent dependent-marking types—emerge rather clearly even from small samples and almost regardless of exemplar-based versus proportion-based definitions. The double-marking and zero-marking types, in contrast, are less frequent overall, and coherent geographical patterns for them emerge only with a sample of comfortably over 200 languages and on exemplar-based definitions.

Furthermore, the distributions of the polar types have emerged clearly in previous work not based on exemplars: the American and Melanesian preponderance of head marking and the Old World preponderance of dependent marking are clear from the smaller sample and nonexemplar definition of Nichols 1986 and 1992 and from Cysouw 2002, where head marking and dependent marking are tracked as independent features. The double-marking and zero-marking types have not been previously mapped (or, more precisely, they are not numerous enough to show up on samples much smaller than this one: see Cysouw 2002: 96–7 for some discussion).

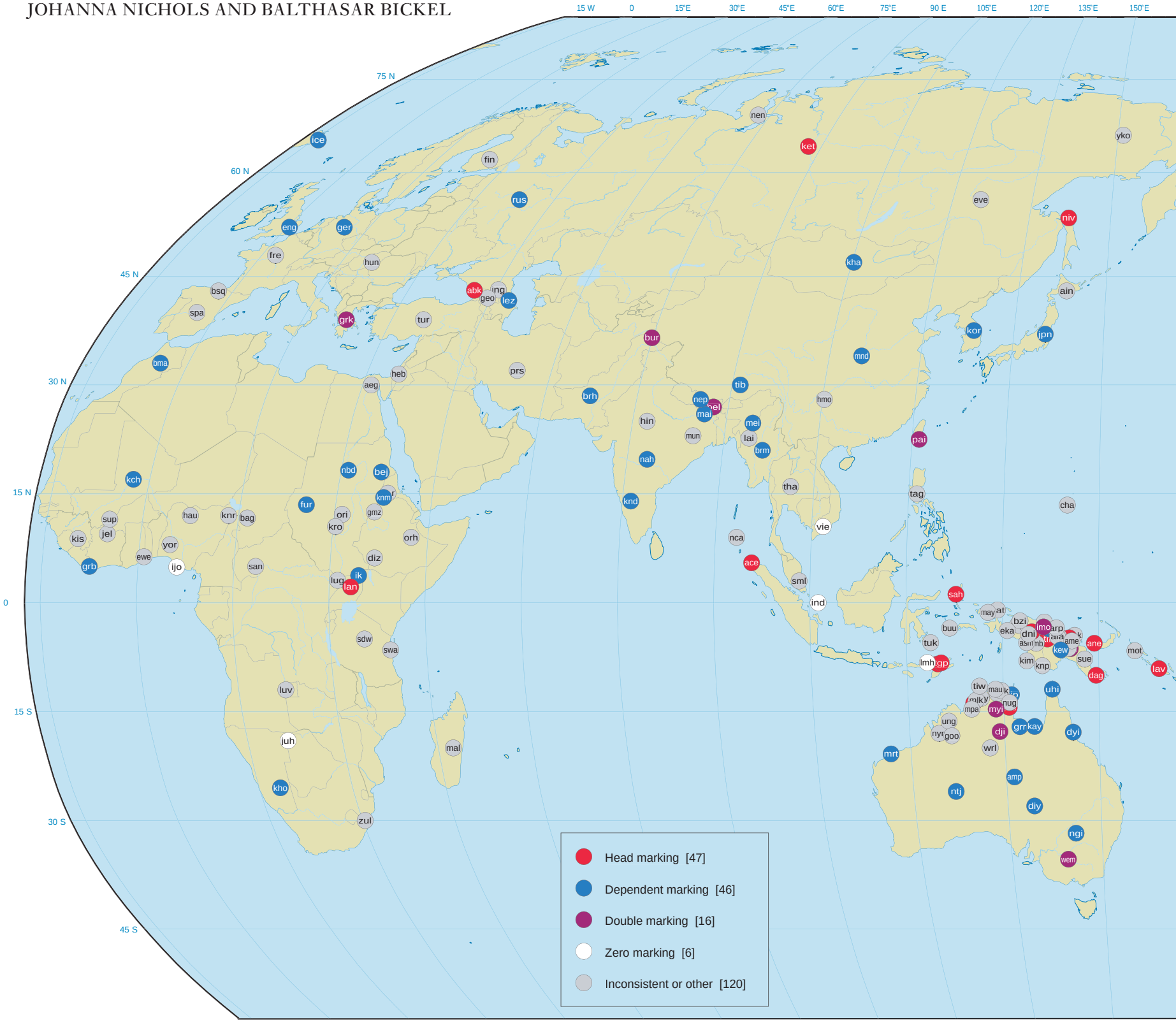
Locus can vary considerably within old families: in these maps, for example, the Uto-Aztecan family includes the consistently head-marking Pipil, spoken in El Salvador, and the consistently dependent-marking Tümpisa Shoshone, spoken in the southwestern United States; the Austronesian family includes the consistently head-marking Acehnese, spoken in Indonesia, and the consistently dependent-marking Maori, spoken in New Zealand. In general, though, locus is of good durability in language families. It appears to be of great durability in areal populations, for the clear Pacific Rim distribution of head marking evidently reflects colonization movements around the Pacific Rim and founder effects in the American linguistic population, and both of these result from population movements that began tens of thousands of years ago. In line with this, we find traces of Pacific Rim locus trends in the Himalayan or Caucasian mountain enclaves that preserve typological profiles antedating the great spreads of interior Eurasian language families. The dependent-marking type is decently attested everywhere but rare in the Americas, where its low frequency must be another result of founder effects. In the Americas, all four types are attested along the Pacific coast but only the head-marking type is common to the east. Whether the diversity of types along the Pacific coast reflects great age for that linguistic population or rather originates in an overlay of recent Eurasian colonizations on an earlier American stratum still remains to be seen.

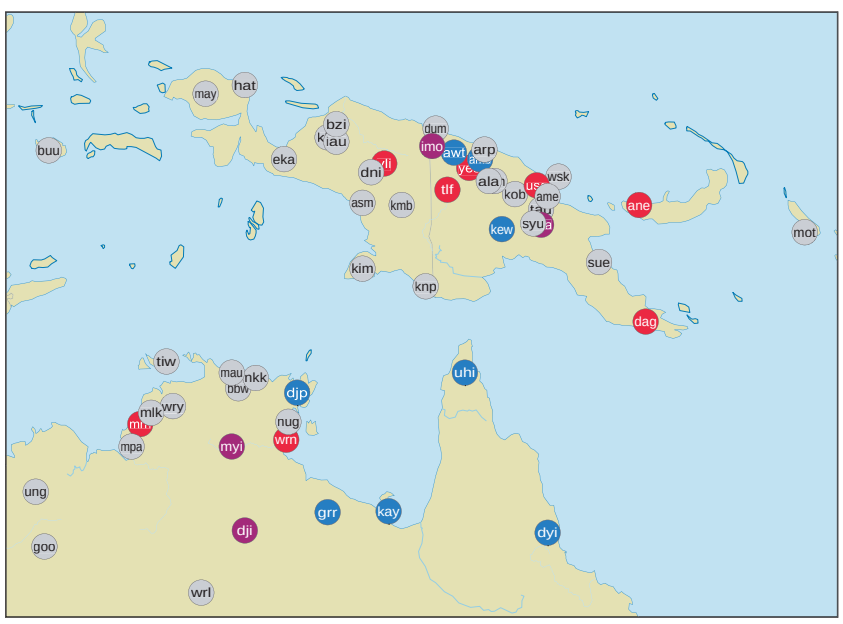
Acknowledgements

This research was supported by US National Science Foundation Grant No. 96-16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P.I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, David A. Peterson, Rebecca Voll, and Suzanne Wilhite for help with data collection.

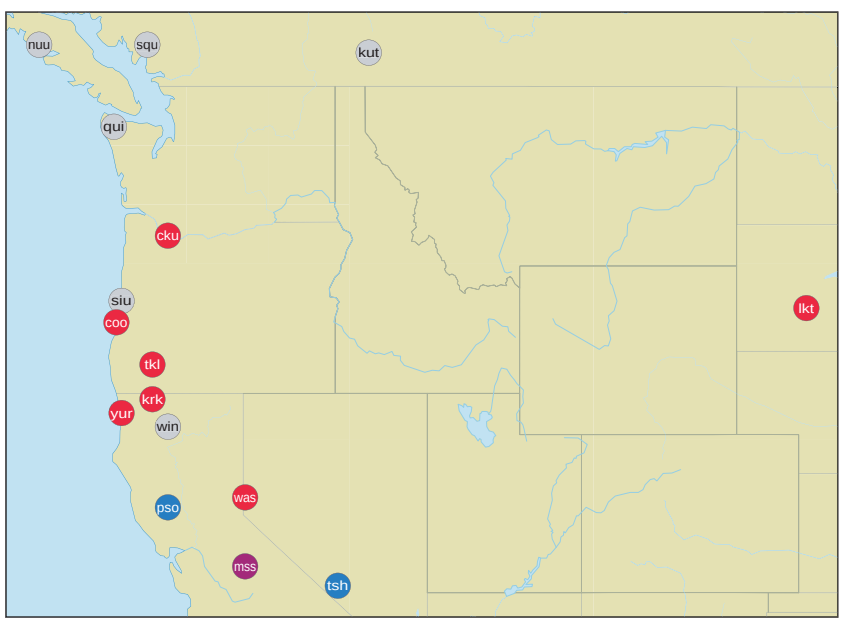
25 Locus of Marking: Whole-language Typology

JOHANNA NICHOLS AND BALTHASAR BICKEL

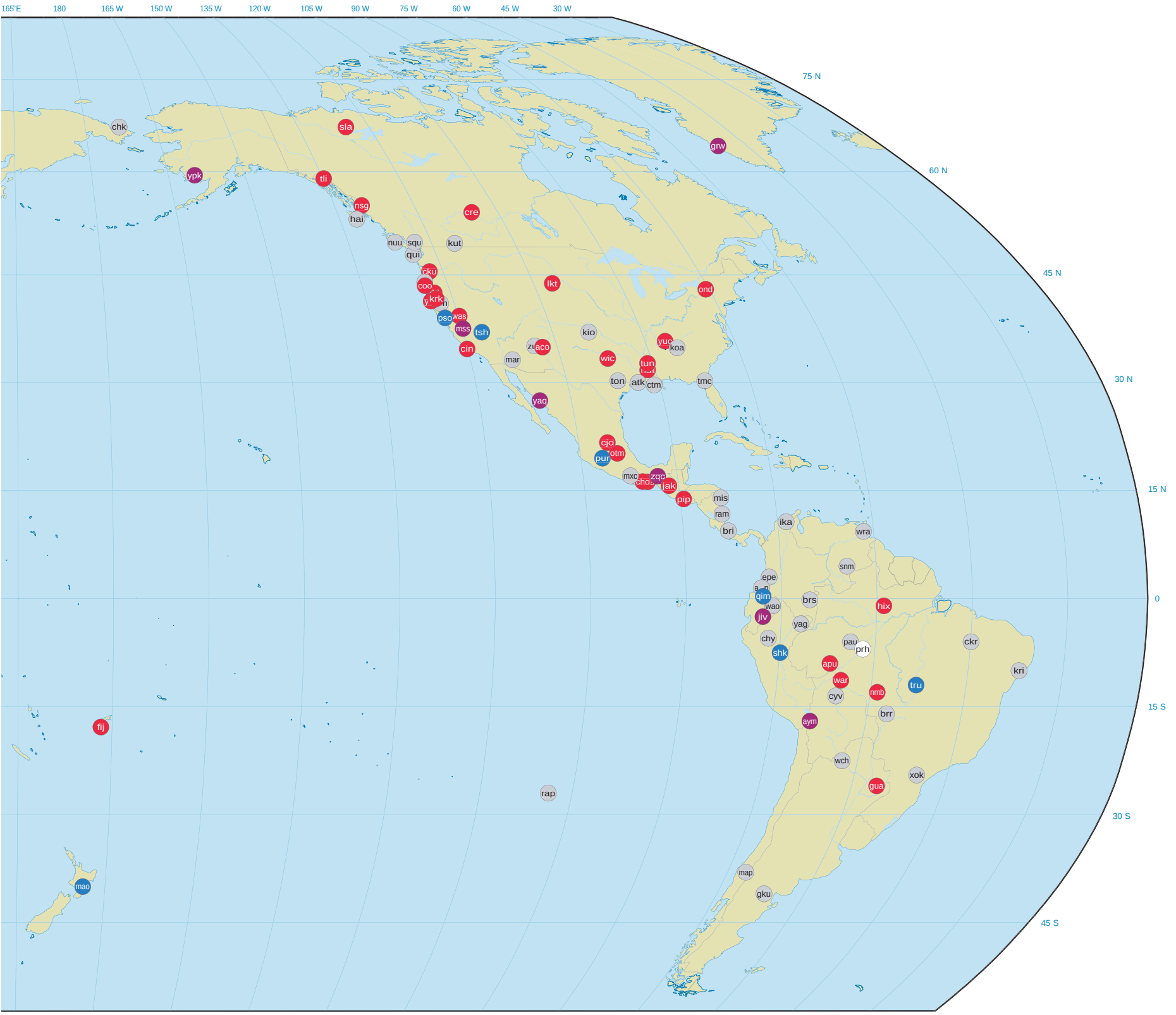


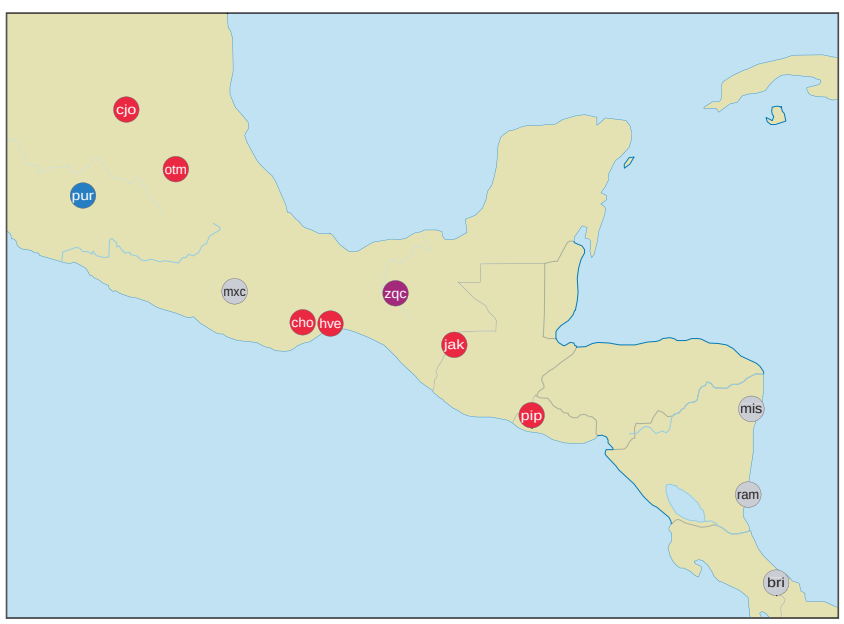


New Guinea and northern Australia

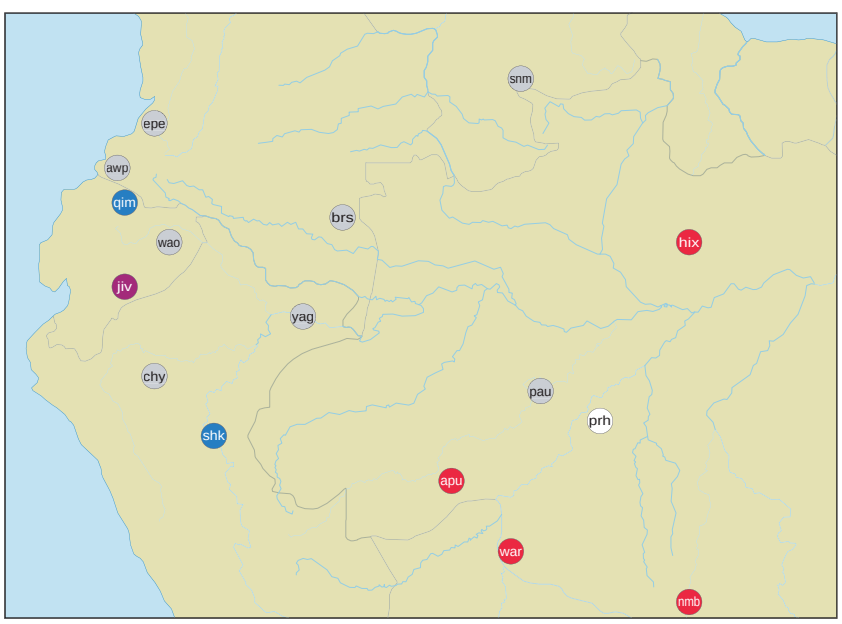


western United States





southern Mexico and Central America



northwestern South America

26 Prefixing versus Suffixing in Inflectional Morphology

MATTHEW S. DRYER

1 Defining the values

This map shows the overall extent to which languages use prefixes versus suffixes in their inflectional morphology. The following inflectional affixes were considered:

(i) case affixes on nouns (see Map 51)
(ii) pronominal subject affixes on verbs
(iii) tense-aspect affixes on verbs (see Map 69)
(iv) plural affixes on nouns (see Map 33)
(v) pronominal possessive affixes on nouns (see Map 57)
(vi) definite or indefinite affixes on nouns (see Maps 37 and 38)
(vii) pronominal object affixes on verbs
(viii) negative affixes on verbs (see Map 112)
(ix) interrogative affixes on verbs (see Map 116)
(x) adverbial subordinator affixes on verbs (see Map 94)

Explanations of these various categories of affixes can be found in the chapters accompanying the maps cited in the above list. Pronominal subject affixes and pronominal object affixes are defined similarly to the definitions used in Chapters 100, 102, and 104 except that (i) affixes coding gender or number but not person are included here, and (ii) clitics that can attach to words other than the verb are not included here.

Languages were assigned a **prefixing index** and a **suffixing index** in the following way. For each affix type above for which the language predominantly employs prefixes, one point was assigned to the prefixing index for the language, and analogously for the suffixing index. If for a given affix type, a language has both prefixes and suffixes with neither deemed dominant, half a point was added to both the prefixing index and the suffixing index of the language. However, the first three affix types in the above list were considered sufficiently important that they were assigned twice as many points as the other affix types: for each of these affix types, if a language predominantly employs prefixes, two points were added to the prefixing index of the language and analogously for suffixes. If for one of these three types, a language has both prefixes and suffixes with neither deemed dominant, one point was added to both the prefixing index and the suffixing index of the language. In the discussion below, I will refer to the sum of the prefixing index and the suffixing index as the **affixing index**. Note that the number of distinct morphemes of a given type was not considered: for example, a language with a single tense-aspect suffix was coded the same as a language with many tense-aspect suffixes. Derivational affixes and other types of inflectional affixes were not considered.

We can illustrate the calculation of these indices for Nuaulu (Austronesian; Seram Island, Indonesia). Nuaulu has pronominal subject prefixes on verbs, as in (1a), contributing two points to its prefixing index; a plural suffix on nouns, also illustrated in (1a), contributing one point to its suffixing index; and both possessive prefixes and possessive suffixes, as in (1b) and (1c), contributing half a point to the prefixing index and half a point to the suffixing index.

(1) Nuaulu (Bolton 1990: 45, 54)

a. *i-au-sipu* *api-a*
3SG-CAUS-get.down thing-PL
'She put down her things.'

b. *au* *we-topi*
1SG 1SG.POSS-hat
'my hat'

c. *unu-e*
head-3SG.POSS
'his head'

Nuaulu has what look like pronominal object suffixes on verbs, but these are excluded here since they attach to a postverbal adverb, when there is one, rather than to the verb, as in (2).

(2) Nuaulu (Bolton 1990: 89)
arei-mo, *wa-ni* *a-rime* *kuru-i*
this-TOP EXIST-PROX 1PL.EXCL-hold tightly-3SG.OBJ
'We are holding it tightly.'

Nuaulu does not have affixes belonging to any of the other categories listed above, so it has a prefixing index of 2.5 and a suffixing index of 1.5. The prefixing index for Nuaulu is 62.5 per cent of the amount of its affixing index (2.5 prefixing points out of a total of 4 = 2.5 + 1.5). According to the definitions given below, Nuaulu counts as a language with a moderate preference for prefixing.

1. Little or no inflectional morphology	122
2. Predominantly suffixing	382
3. Moderate preference for suffixing	114
4. Approximately equal amounts of suffixing and prefixing	130
5. Moderate preference for prefixing	92
6. Predominantly prefixing	54
total	894

The first value shown on the map is for **languages which have little or no inflectional prefixing or suffixing**. A language is classified as a language of this type if its affixing index is two or less. An example of a language of this type is Thai, which is completely lacking in inflectional affixes of the categories examined. A more borderline case of this type is Vai (Mande; Liberia; Welmers 1976), in which the only inflectional affixes I record are suffixes for tense-aspect, which gives the language an affixing index of only two. Other less frequent inflectional methods like infixation, tonal affixes, and stem changes were ignored, so that a language might count as a language with little inflectional prefixing or suffixing but still have affixation of these other types. For example, Dinka (Nilotic; Sudan; Nebel 1948) employs stem changes for case and for plural, but the only suffixes or prefixes I record are a definite suffix and possessive suffixes, which give the language an affixing index of only two, which means that it is shown as having little or no inflectional prefixing or suffixing.

For all the remaining types, the affixing index must be greater than two. The types differ from each other in the relative amount of prefixing and suffixing.

The second type is **languages which are predominantly suffixing**, defined for the purposes of this map as languages with a suffixing index which is more than 80 per cent of its affixing index. The highest suffixing index in the sample is eleven, represented by two languages, West Greenlandic (Eskimo; Fortescue 1984) and Central Yup'ik (Eskimo; Alaska; Reed et al. 1977); both of these languages are exclusively suffixing for the affix categories examined. This type also includes languages with considerably less affixation, but what affixation they have is largely if not entirely suffixing, as long as the affixing index is greater than two. For example, Korana (Central Khoisan; South Africa; Meinhof 1930) has an affixing index of three, with suffixes for case (two points) and plural (one point) and no inflectional prefixes.

The third type is **languages with a moderate preference for suffixes**, defined as languages in which the suffixing index is more than 60 per cent of the affixing index but not more than 80 per cent. An example of such a language is Beja (Cushitic; Sudan; Reinisch 1893), which has a suffixing index of ten and a prefixing index of three (so that its suffixing index is 77 per cent of its affixing index). An example of a language of this type with less morphology is Mokilese (Oceanic; Micronesia; Harrison and Albert 1976), which has a suffixing index of two and a prefixing index of one.

The fourth type is **languages with approximately equal amounts of suffixing and prefixing**, defined here as languages with a suffixing index that is greater than or equal to 40 per cent of the affixing index and less than or equal to 60 per cent of the affixing index. An example of a language of this type with considerable inflectional morphology is Ubykh (North-West Caucasian; Turkey; Charachidze 1989), whose suffixing index and prefixing index are both 5.5. An example of a language of this type with less inflectional morphology is Kiribati (Oceanic; Kiribati; Groves et al. 1985), whose suffixing index and prefixing index are both two.

The fifth type is **languages with a moderate preference for prefixes**, where the prefixing index is more than 60 per cent of the affixing index but not more than 80 per cent. An example of such a language is Mohawk (Iroquoian; New York State and Ontario; Bonvillain 1973), which has a prefixing index of six and a suffixing index of three. An example of this type with less morphology is Au (Torricelli; Papua New Guinea; Scorza 1985), which has a prefixing index of two and a suffixing index of one. Nuaulu, used above to illustrate the calculation of the indices, is also a language of this type.

The last type is **languages which are predominantly prefixing** in their inflectional morphology, defined here as languages with a prefixing index that is more than 80 per cent of its affixing index. The highest prefixing index in the sample is 9.5 and is found in Hunde (Bantu; Democratic Republic of Congo; Kahombo 1992). Hunde is not exclusively prefixing; it has a suffixing index of 0.5, due to its having both possessive prefixes and possessive suffixes, with neither dominant. Again, this type includes languages with less morphology, as long as their affixing index is greater than two and their affixes are primarily prefixes. For example, Sango (Adamawa-Ubangi, Niger-Congo; Central African Republic; Samarin 1967) has an affixing index of three, with pronominal prefixes on verbs (two points) and plural prefixes on nouns (one point).

2 Geographical distribution

The map shows an overall preference for suffixes. In fact, of the five types other than those with little affixation, those which are predominantly suffixing are about as frequent as the other four types combined (382 versus 390), and outnumber those which are predominantly prefixing by about seven to one. In general, the areas where one finds languages with a moderate preference for suffixes are similar to those where one finds a strong preference for suffixes, and similarly to a lesser extent for prefixes, so it is possible to speak of two basic types, languages with more prefixing and languages with more suffixing, and to discuss the geographical distribution of these two basic types.

The preference for suffixes is especially strong on the mainland of Eurasia, where there are very few languages with more prefixing: in fact only five exceptions are shown, Ket (Yeniseian; Siberia; Werner 1997), Temiar (Aslian, Mon-Khmer; Malaysia; Benjamin 1976), and three Tibeto-Burman languages. Even languages with approximately the same amount of prefixing and suffixing are quite infrequent in this area. The suffixing preference is even stronger in Australia, where languages with a prefixing preference are restricted to the north; the languages in most of Australia not only display a suffixing preference, but a strong one. Languages with more suffixing also form the dominant type in South America, though for many of the languages here, the preference is a weak one. New Guinea also exhibits a suffixing preference, at least in areas closer to the centre. Around the edges, there are some languages with a preference for prefixes; some of these are Austronesian languages, but others belong to the Torricelli, East Bird's Head, and West Papuan families.

In North America, there are more languages preferring suffixes overall, but there are also clear areal patterns. Within 800 kilometers of the western coast of Canada and the United States, the languages are predominantly suffixing. To the east of this area, the two basic types are about equally common. Languages with more prefixing are in the majority in Mesoamerica.

In Africa, both types are common overall, again with clear areal patterns. In the southern half of the continent, languages with more prefixes predominate: the majority of these are Bantu languages, but one also finds more prefixing among the Adamawa-Ubangi languages to the north of Bantu. The greatest number of languages on the map that are predominantly prefixing (and not just moderately prefixing) is in Africa. Although the majority of these are Bantu languages, there are also a number of other Niger-Congo and Nilo-Saharan languages.

Languages with approximately equal amounts of prefixing and suffixing are more common in areas where there are both many languages with more prefixing and many languages with more suffixing, notably in central Africa and North America, rather than in areas where languages with suffixes predominate. Such languages are also common in South America.

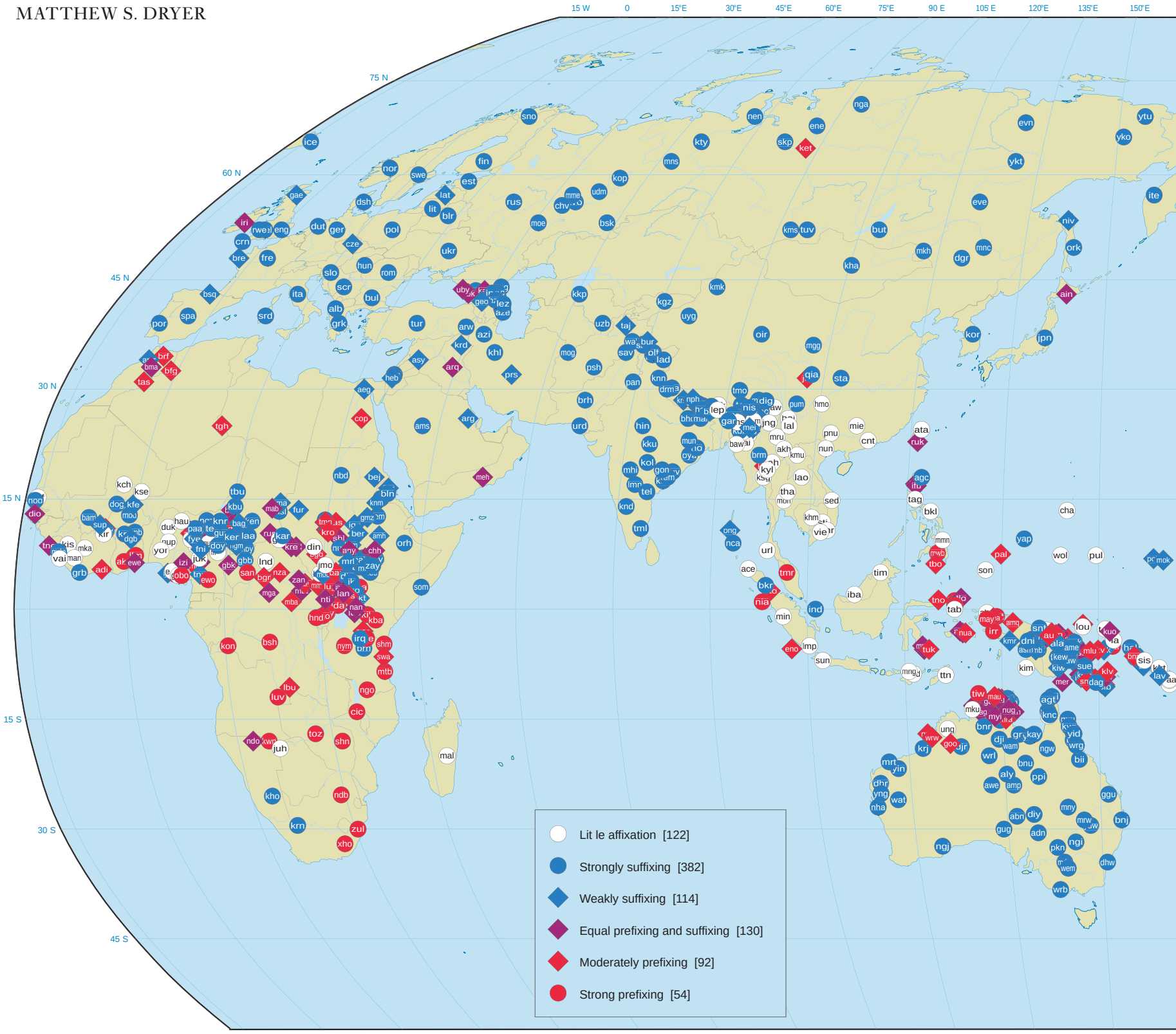
Languages with little inflectional prefixation or suffixation are concentrated in two areas. The larger area stretches from South-East Asia south to the Austronesian languages of Indonesia, the Philippines, and east into the Pacific. The other area is in a belt across the middle of Africa, though there are also many languages in this region with somewhat more inflectional morphology. But the only area in which all languages shown are of this type is on mainland South-East Asia, from Vietnam to Thailand and stretching north into adjacent areas in China. In the other areas where languages of this type are common, among Austronesian languages and in Africa, they are interspersed with languages of other types. Among Austronesian languages, those other types are typically ones involving more prefixing. In the area in Africa where languages with little or no affixation are common, all other types are common as well.

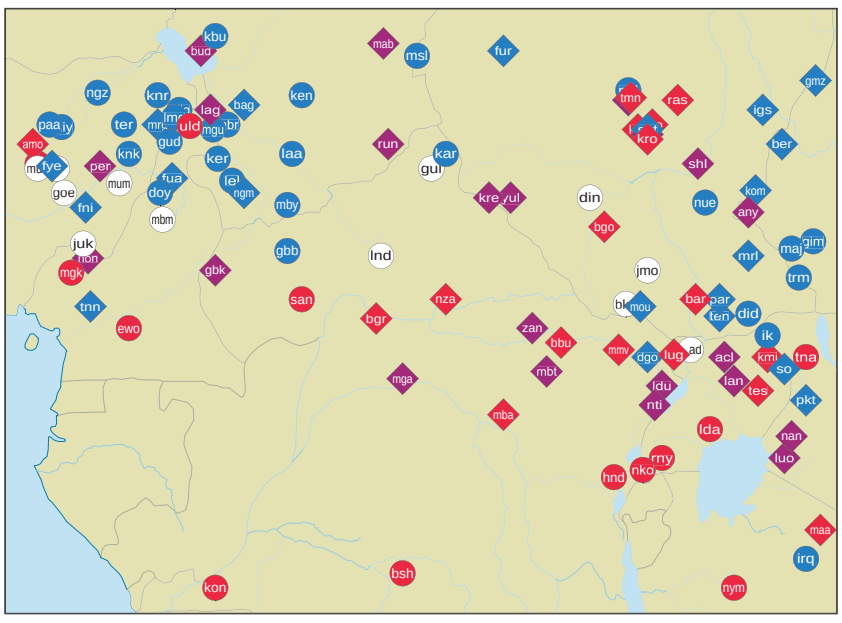
3 Theoretical issues

Perhaps the largest theoretical question is why suffixes are more frequent than prefixes. Various hypotheses have been offered. Among them is the idea that prefixes make lexical recognition more difficult, especially if it is more difficult to identify the beginning of stems (Cutler et al. 1986). Suffixes do not present a problem, since identifying the ends of stems is less important for lexical recognition. Further discussion is found in Greenberg (1957), Hall (1988), and Bybee et al. (1990). It should be noted that different categories of affixes exhibit different degrees of preference for suffixes. For example, case affixes exhibit a particularly strong suffixing preference; case prefixes are fairly rare (see Map 51). On the other hand, pronominal possessive prefixes are approximately as common as suffixes (see Map 57).

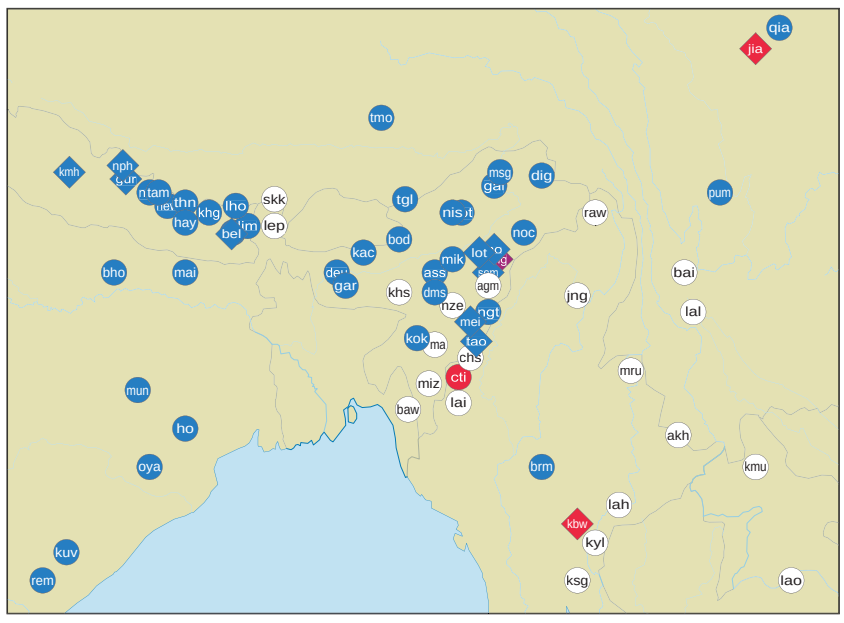
26 Prefixing versus Suffixing in Inflectional Morphology

MATTHEW S. DRYER

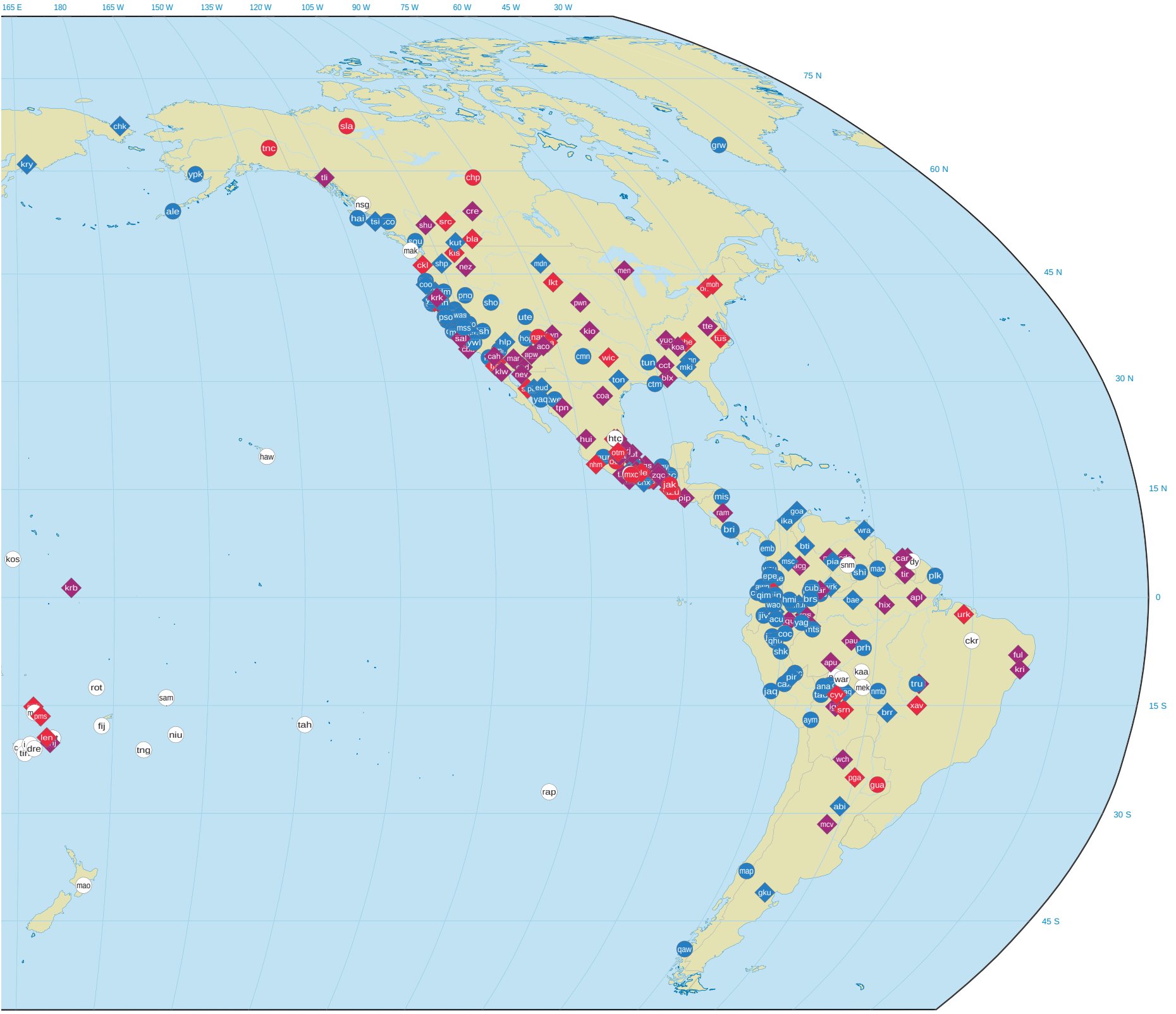


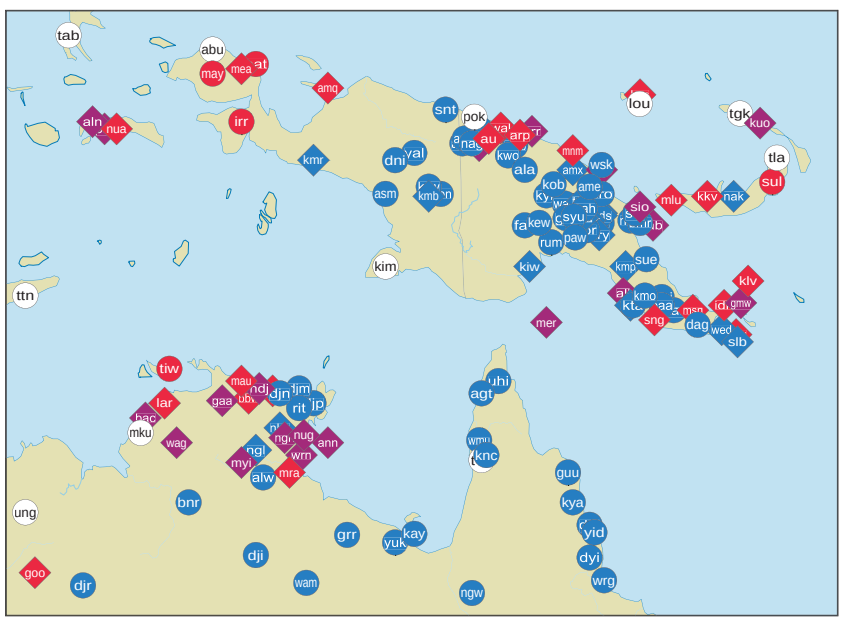


central Africa

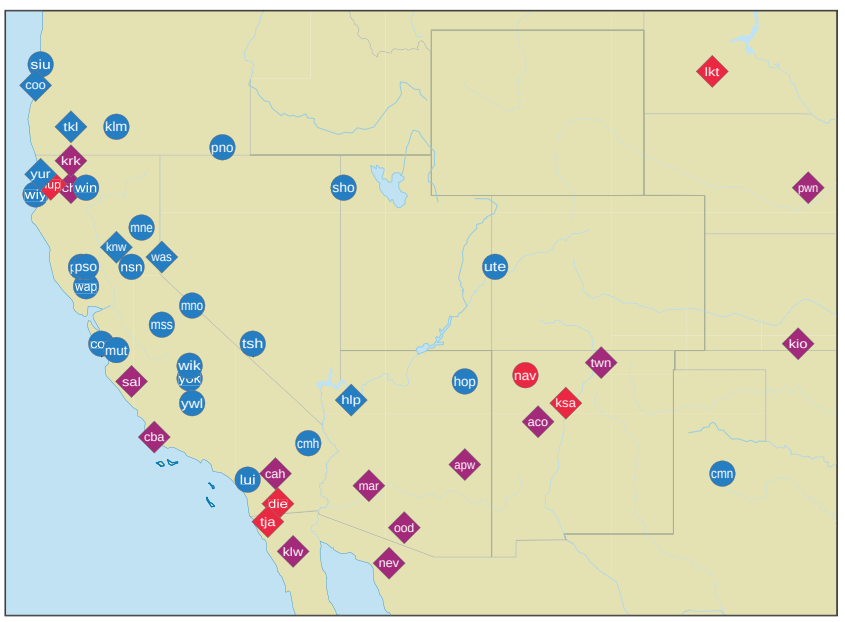


Nepal, eastern India, and surrounding areas





New Guinea and northern Australia



western United States

27 Reduplication

CARL RUBINO

1 Features Values

The repetition of phonological material within a word for semantic or grammatical purposes is known as **reduplication**, a widely used morphological device in a number of the world's languages. The languages classified on the accompanying map are sorted into three categories: languages that do not employ reduplication as a grammatical device, languages that productively employ both partial and full reduplication, and languages that only employ full reduplication.

Full reduplication is the repetition of an entire word, word stem (root with one or more affixes), or root. Examples are Nez Perce (Sahaptian; northwestern United States) full word lexical reduplication: *té:mul* 'hail' versus *temulté:mul* 'sleet' (Aoki 1963: 43), or Tagalog full root reduplication, shown here with the verbalizing prefix *mag-*, where the reduplicant *isip* is identical to the base *isip* 'think': *mag-isip* 'to think' versus *mag-isip-isip* 'to think about seriously'.

Partial reduplication may come in a variety of forms, from simple consonant gemination or vowel lengthening to a nearly complete copy of a base. In Pangasinan (Austronesian; Philippines) various forms of reduplication are used to form plural nouns.

(1) Pangasinan (Rubino 2001: 540)

CV-	*toó*	'man'	*totóo*	'people'
-CV-	*amigo*	'friend'	*amimígo*	'friends'
CVC-	*báley*	'town'	*balbáley*	'towns'
C_1V-	*plato*	'plate'	*papláto*	'plates'
CVCV-	*manók*	'chicken'	*manómanók*	'chickens'
Ce-	*duég*	'carabao'	*deréweg*	'carabaos'

The following feature values are represented on the map:

1.	Productive full and partial reduplication	277
2.	Full reduplication only	34
3.	No productive reduplication	56
	total	367

It has been observed that languages with productive partial reduplication will most likely also make use of full reduplication (Moravcsik 1978: 328), making semantic and grammatical distinctions in the use of the two reduplicative types, as seen in Squamish (Salishan; British Columbia).

(2) Squamish (Kuipers 1967: 99–100)

total:	*kʷaɨ̓ʔkʷaɨ̓*	'play hide and seek'
		from the root *kʷaɨ̓*
partial:	*s-λ'lλ'lmut*	'old people'
		from the singular *s-λ'lmut*

Languages that employ partial reduplication may do so in various ways. Reduplicated material is most often found at the beginning of a base, but occurs also in medial and final position.

(3) Reduplicative prefixes, suffixes and infixes:

a. Hunzib (Nakh-Daghestanian; eastern Caucasus)
initial CV(C) reduplication (van den Berg 1995: 34)

bat'iyab	'different'	*bat'bat'iyab*	'very different'
muğáƛ	'after'	*mumuğáƛ*	'much later'

b. Choctaw (Muskogean; Mississippi and Alabama)
medial CV reduplication (Kimball 1988: 440)

tonoli	'to roll'	*tononoli*	'to roll back and forth'
binili	'to sit'	*bininili*	'to rise up and sit down'

c. Paumarí (Arauan; Amazonas, Brazil)
final disyllabic reduplication (Chapman and Derbyshire 1991)
a-odora-dora-bakhia-loamani-hi
1PL-gather.up-REDUP-frequently-really-THEME
'We keep gathering them.'

The phonological nature of the reduplicated material varies from language to language and construction to construction. Reduplicative morphemes can be characterized by number of phonemes included in the copy, C, CV, CVC, V, CVCV, etc. (see 1, 3); the number of syllables to be reduplicated (see 3); or the number of repeated morae. In Ngiyambaa (Pama-Nyungan; New South Wales, Australia), the reduplicant consists of a copy of the first syllable and a copy of a light version of the second syllable, not including final vowel lengthening or a coda consonant (Donaldson 1980): *magu-magu*: 'around one,' *dhala-dhalarbi-ya* [REDUP-shine-PRES] 'to be pretty shiny'. In some cases, even the number of times a sequence is reduplicated is a morphological factor, e.g. Mokilese duplication versus triplication: *roar* 'give a shudder' > *roarroar* 'be shuddering' > *roarroarroar* 'continue to shudder' (Harrison 1973).

Reduplicative constructions can also be characterized as being simple, complex, or automatic. A **simple** construction is one in which the reduplicant matches the base from which it is copied without phoneme changes or additions (see 3). A **complex** construction involves reduplication with some different phonological material, such as a vowel or consonant change or addition, or morpheme order reversal. Mangarrayi (Northern Territory, Australia) has a reduplicative construction to denote plurality in which the consonant of the second syllable and the vowel of the first syllable are copied to form a new second syllable in the derived word. The newly created syllable does not correspond to any constituent in the original word: *walima* 'young person' > *walalima* 'young people'; *yirag* 'father' > *yirirag-ji* 'father and children' (Merlan 1982). In Tuvan (Turkic; Russia), diminutive '*s*' reduplication copies the entire base except the initial consonant, which is replaced by [s] in the reduplicant, e.g. *pelek* 'gift' > *pelek-selek* 'gift.DIMINUTIVE'. For bases that are vowel-initial, an onset [s] is added to the reduplicant, e.g. *aar* 'heavy' > *aar-saar* 'heavy.DIMINUTIVE'; *uuruk-suuruk* 'simultaneously' (Harrison 2000). Patterns such as these exist in a number of languages and are collectively referred to as "echo constructions".

Nias (Sundic, Austronesian; Sumatra) disyllabic reduplication sometimes includes voicing: *a-fusi* 'white'; *a-vuzi-vuzi* 'whitish' (Brown 2001). Reduplication can also be discontinuous, in which case a small segment is inserted between the reduplicant and the base. In Alamblak (Sepik; Papua New Guinea), *ba* joins reduplicated constituents in an intensifying construction:

(4) Alamblak (Bruce 1984)
hingna-marŋa-ba-marŋa-më-r
work-REDUP-*ba*-straight-REMOTE.PST-3SG.M
'he worked very well'

Automatic reduplication is reduplication that is obligatory in combination with another affix, and which does not add meaning to the overall construction; the affix and reduplicated matter together are monomorphemic, e.g. the Ilocano (Austronesian; Philippines) pretentative prefix *aginCV-*: *singpet* 'behave', *agin-si-singpet* 'to pretend to behave'.

2 Function

Reduplicative morphemes can carry a number of meanings, and in some languages the same reduplicative morpheme is used to

denote quite contrary meanings. For example, the Ilocano CVC-distributive prefix for nouns, when applied to numbers, specifies limitation: *sab-sábong* 'various flowers', *wal-waló* 'only eight'. With verbs (and adjectives), reduplication may be used to denote a variety of things, such as number (plurality, distribution, collectivity), distribution of an argument, tense, aspect (continued or repeated occurrence, completion, inchoativity), attenuation, intensity, transitivity (valence, object defocusing), or reciprocity. For example, Alabama (Muskogean; Alabama) marks the temporary versus permanent distinction in verbal aspect via reduplication (vowel lengthening): *loca* 'to be black (covered in soot)' versus *lóoca* 'to be a black person', as well as attenuation via gemination: *kasatka* 'cold' > *kássatka* 'cool', *lamatki* 'straight' > *lámmatki* 'pretty straight' (Hardy and Montler 1988). Luiseño (Uto-Aztecan; California) employs two types of reduplication quite iconically to denote various plural actions: *lawi* 'to make a hole', *law-lawi* 'to make two holes, make a hole twice', *lawa-láwi* 'to make many holes, more than two' (Kroeber and Grace 1960). Similarly, Lampung (Sundic, Austronesian; Sumatra) uses different reduplicative constructions to signal varying degrees of intensity: *balak-balak* 'very large', *xa-xabay* 'somewhat afraid' (Walker 1976). Arapesh (Torricelli; Papua New Guinea) employs reduplication to intensify or distribute the meaning of an action, often implying carelessness or lack of control on the part of the agent: *su* 'touch, hold', *susu* 'touch all over, paw'; *ɾ̥ipok* 'cut', *ɾ̥iɾ̥ipok* 'hack up' (Dobrin 2001: 36).

With nouns, reduplicative morphemes have been known to denote concepts such as number (see 1), case, distributivity, indefiniteness, reciprocity, size (diminutive or augmentative), and associative qualities. For instance, Ilocano reciprocals: *balem-bales* (CVCN-revenge) 'avenge each other' (Rubino 2000: 84); Chukchi (Chukotko-Kamchatkan) absolutive singulars: *jokwa-t* (eider.duck-PL) 'eider ducks' versus *jokwa-jow* (eider.duck-REDUP[ABS.SG]) 'eider duck, absolutive' (Dunn 1999: 132); and Yawelmani (Penutian; California) associatives *k'ɔhis* 'buttocks' > *k'ɔk'ɔhis* 'one with large buttocks' (Newman 1944).

With numerals, reduplication has been found to express various categories including collectives, distributives, multiplicatives, and limitatives. For example, Santali (Munda; India) *gɛ-gɛl* 'ten each, by tens'; Pangasinan limitatives *tal-talora* 'only three'; Ao (Tibeto-Burman; India) final CVC reduplication distributives *asem* 'three' > *asemsem* 'three each', *ténet* 'seven' > *ténetnet* 'seven each' (Gowda 1975: 39). (See also Chapter 54 on distributive numerals.)

Reduplication is also used derivationally to alter word class, e.g. Kayardild (Tangkic; Queensland, Australia) *kandu* 'blood' > *kandukandu* 'red' (Evans 1995); Luiseño (Uto-Aztecan) *lepi* 'to tan, soften' > *lepé-lpi-š* 'pliable' (Kroeber and Grace 1960); Tigak (Austronesian) *giak* 'send' > *gigiak* 'messenger' (Beaumont 1979); Khoekhoe (Central Khoisan; Namibia) causatives *!óḿ* 'difficult' > *!óḿ!om* 'make something difficult' (Hagman 1977: 18; note that the tone of the second syllable is lowered to mid tone).

Languages on the map are classified as having a productive reduplicative morpheme only if the morpheme can be systematically generalized to a set of open-class words, and the morpheme can still be applied in the modern form of the language. Modern Greek, for example, is classified as a language that does not meaningfully employ reduplication, although there are a few reduplicative forms present in the modern language that are remnants of a previously productive reduplicative process. In Ancient Greek, the perfect was formed by a *Ce-* reduplicative prefix, e.g. *gé-grapha* 'I have written'; the modern equivalent is now periphrastic *éxo ɣráp-si* [I.have write-PTCP]. The old construction still appears, however, in some learned words, e.g. *ðe-ðo-ména* (Ce-give-MEDIOPASSIVE) 'data', *je-ɣon-os* (Ce-become/happen-PERFECT) 'event'. Modern Greek has also borrowed from Turkish a nonproductive reduplicative prefix used with at least one affective/intensive adjective: *tsir-tsiplákis* 'buck naked' from *tsiplákis* 'naked' (cf. Turkish *bem-beyaz* 'very white' from *beyaz* 'white').

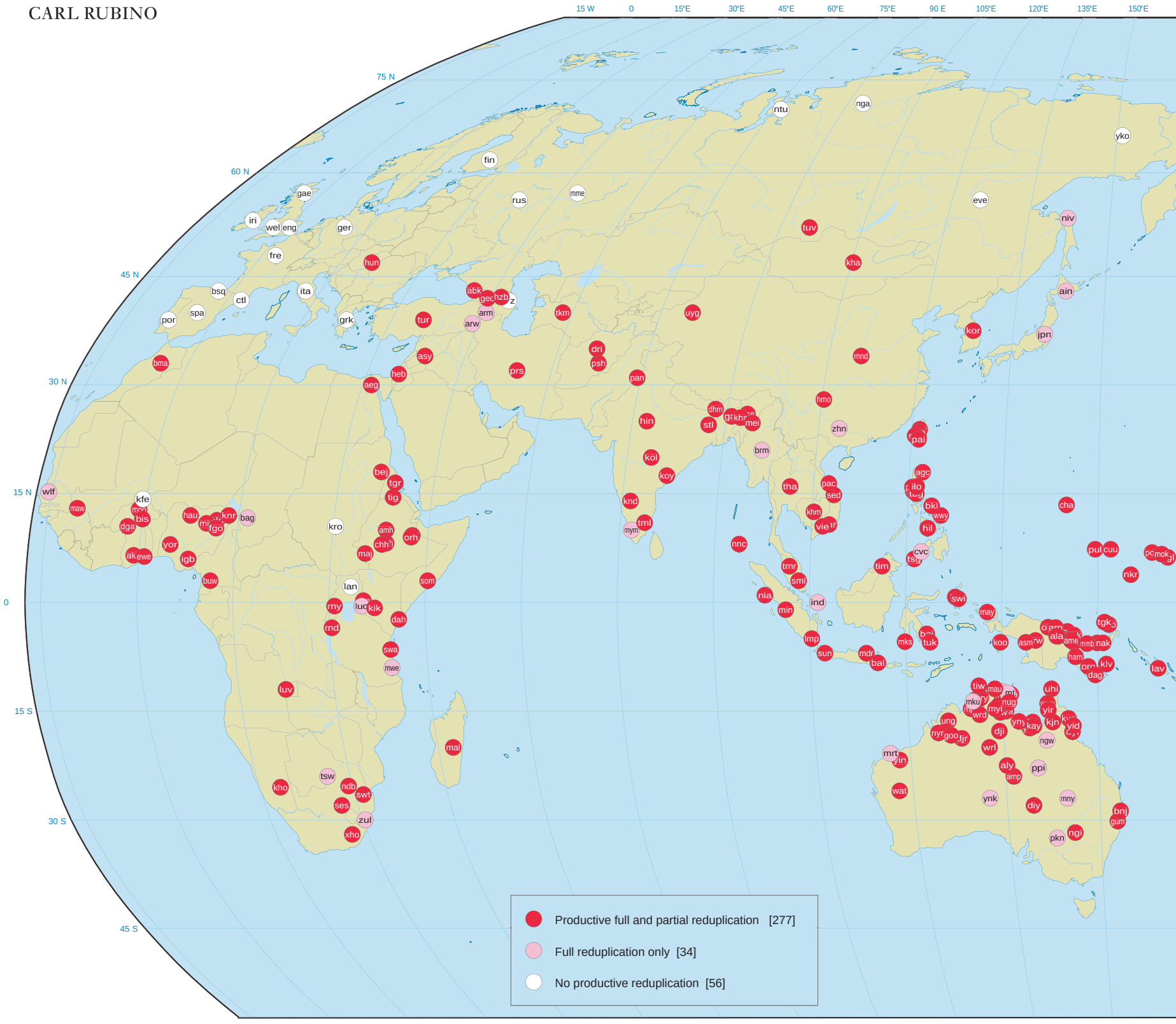
As can be seen from the map, reduplication is a much more pervasive phenomenon than someone coming from a Western European world view might imagine. Reduplication is very common throughout Austronesian (Pacific islands, Philippines, Indonesia, Madagascar), Australia, South Asia, and many parts of Africa, the Caucasus, and Amazonia. In the western hemisphere, some language families are particularly amenable to reduplication (Salishan, Pomoan, Uto-Aztecan, Algonquian, Yuman, Sahaptian, Siouan, etc.) while others are not, such as Athabaskan and Eskimo-Aleut. Western Europe is one area where reduplication does not play a critical role in the morphology. However, many Indo-European languages in the east, which are in contact with other language families, do have reduplicative morphemes.

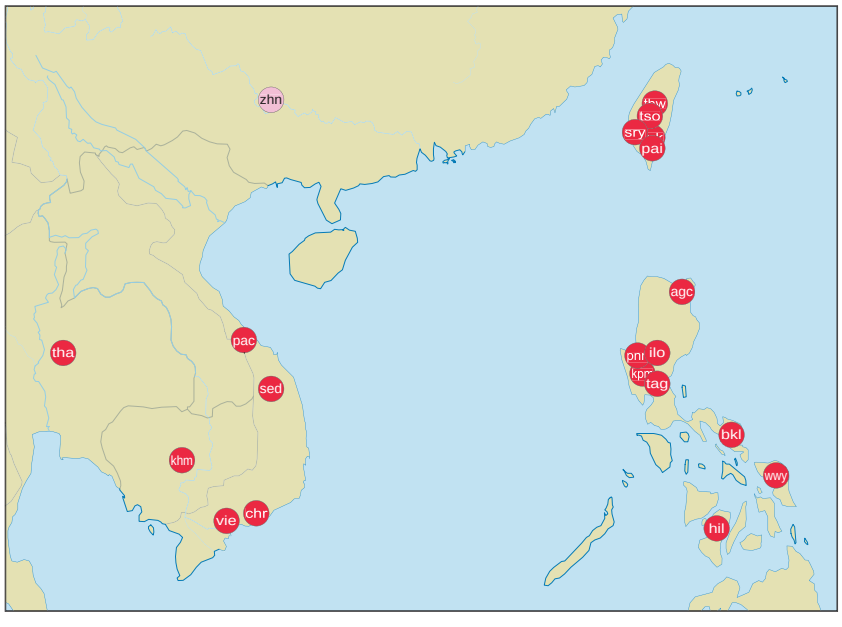
Acknowledgement

Special thanks to Brian Joseph for the Greek data.

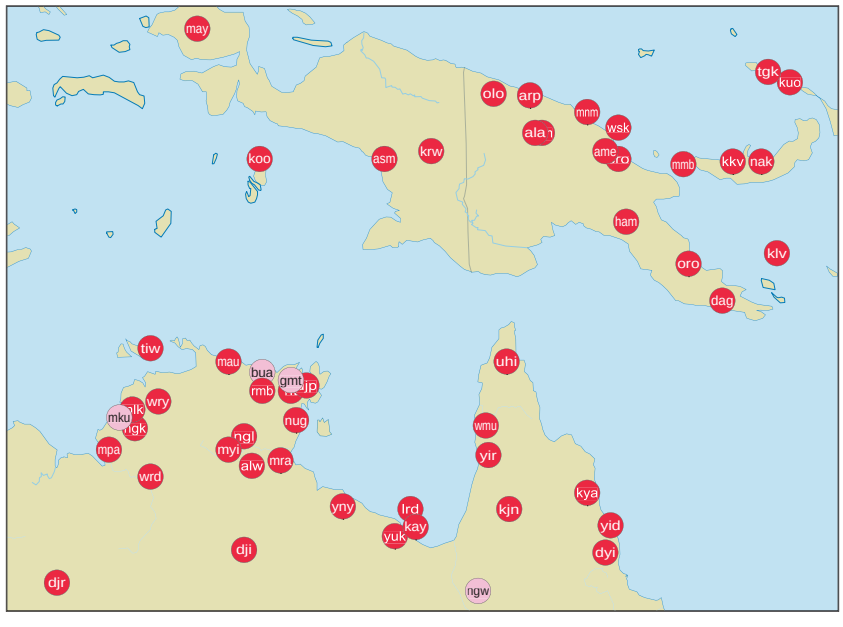
27 Reduplication

CARL RUBINO

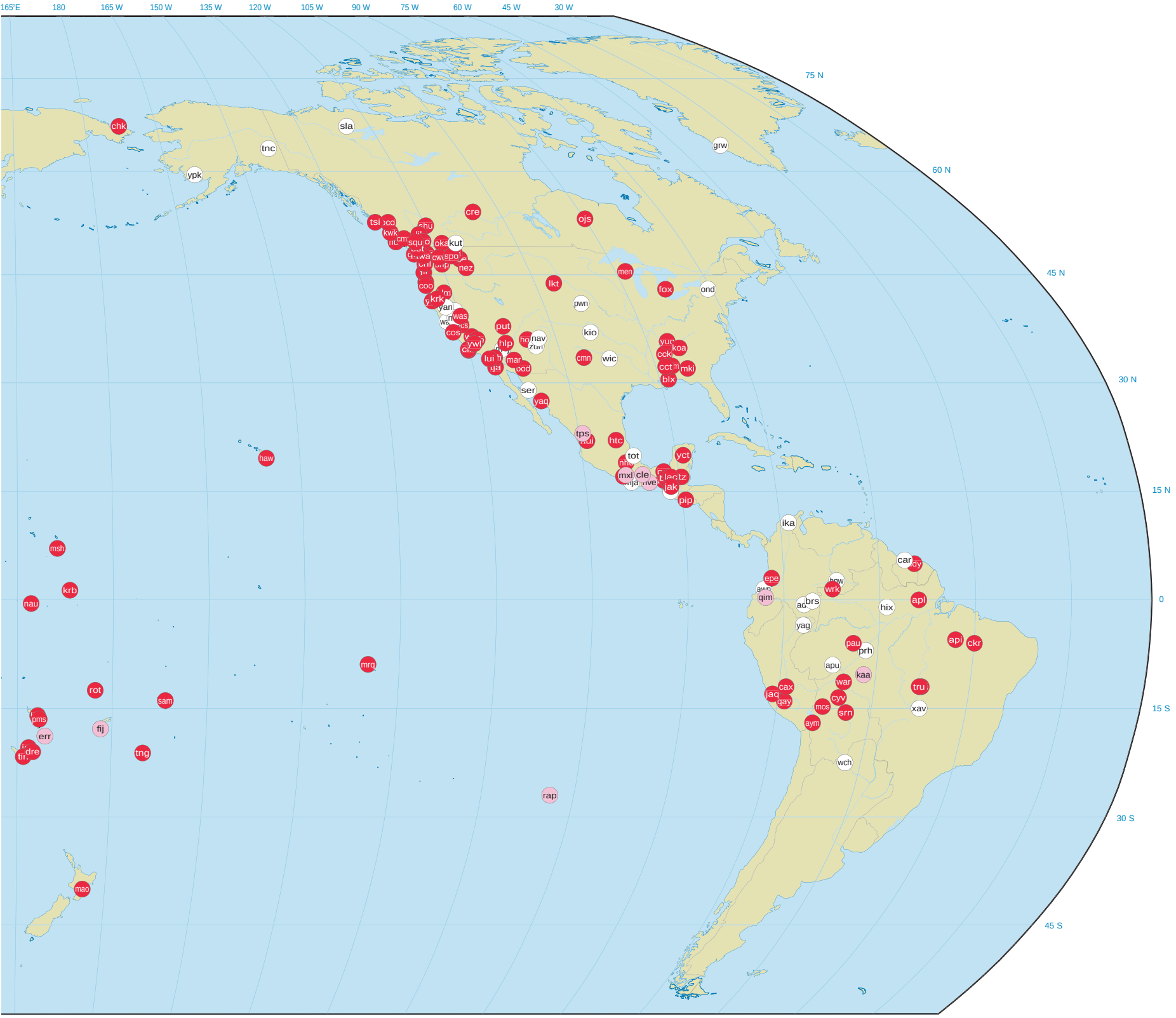


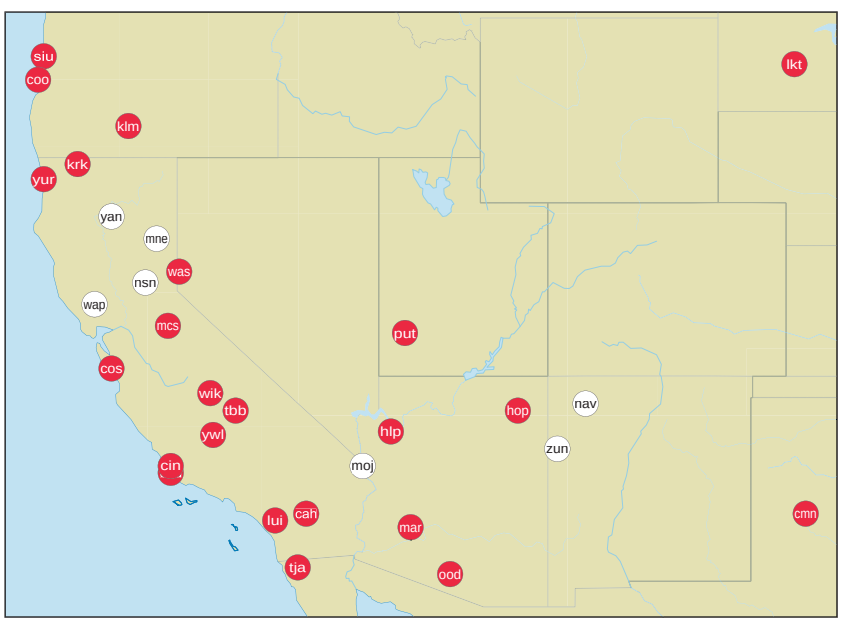


Indochina, Philippines, and adjacent areas

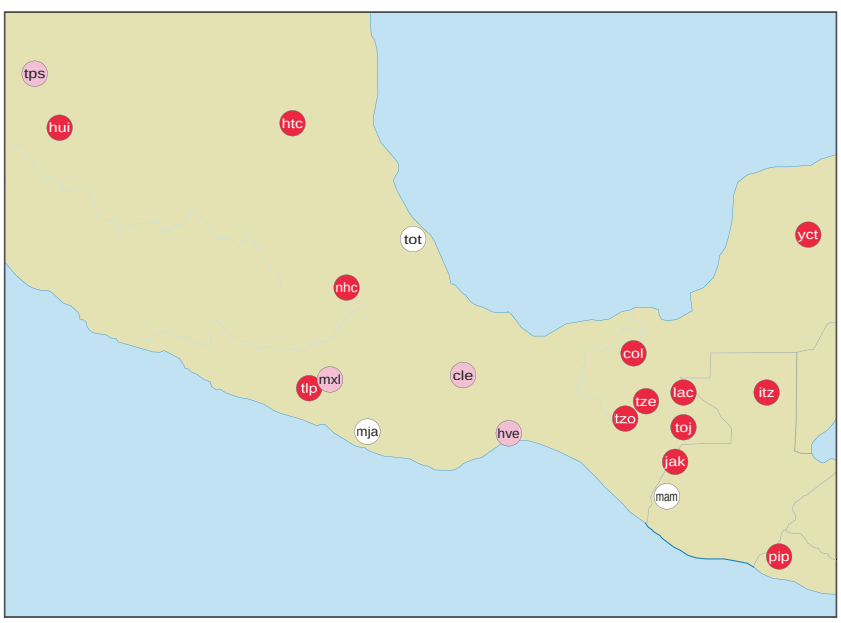


New Guinea and northern Australia





western United States



Mesoamerica

28 Case Syncretism

MATTHEW BAERMAN AND DUNSTAN BROWN

1 Identifying case syncretism

This map shows instances of case syncretism in nominals (nouns, pronouns, and adjectives). We identify case syncretism when a single **inflected form** corresponds to two or more **case functions**. The criterion used here for identifying case functions is form-based: a distinct morphosyntactic case is recognized for a language if it is ever correlated with a formal inflectional distinction (Comrie 1991: 44–7). For example, in Krongo, subject and object forms are never distinguished, so this is not construed as an example of case syncretism, but rather as a single nominative-accusative case (see (1)). However, the dative case is distinct in nouns but identical to this nominative-accusative form in pronouns, so a dative = nominative-accusative syncretism is recognized for Krongo. The formal criteria for identifying syncretism may also be found within the paradigm of a single word, as in Central Yup'ik, where the absolutive and relative (a case combining the functions of ergative and genitive) are distinct in the singular but not in the plural or dual (see (2)).

(1) Krongo (Kadugli; Sudan; Reh 1985: 144–54, 164)

	'I'	'person'
NOM–ACC	*àʔàŋ*	*káaw*
DAT	*àʔàŋ*	*àkáaw*

(2) Central Yup'ik (Eskimo; Alaska; Jacobson 1995: 469–70)

'land'	PL	DU	SG
ABS	*nunat*	*nunak*	*nuna*
REL	*nunat*	*nunak*	*nunam*

2 Defining the values

The map records four values:

1.	Inflectional case marking is absent or minimal	123
2.	Inflectional case marking is syncretic for core cases only	18
3.	Inflectional case marking is syncretic for core and non-core cases	22
4.	Inflectional case marking is never syncretic	34
	total	197

A language is classed as having minimal case marking if the nominal paradigm contains no more than two forms. In such a language the absence of any case distinction is equivalent to the complete absence of case inflection. Since we construe syncretism as a phenomenon affecting inflected words, such languages are here given the same value as languages with no inflectional case marking; i.e. the conditions for identifying syncretism are not met. The **core cases** referred to by value two and value three (cf. Blake 1994: 119–44) prototypically represent the functions of subject and object, and are familiar under the names *nominative* (subject), *accusative* (object), *ergative* (subject of a transitive verb), and *absolutive* (subject of an intransitive verb and object of a transitive). Other cases are classed as non-core cases.

2.1 Syncretism of the core cases. Value two involves the collapse solely of the core cases, i.e. nominative and accusative (see (3)) or ergative and absolutive (see (4)). Some languages display both patterns, resulting in so-called split ergativity, where one set of nominals appears to decline according to a nominative-accusative pattern, another according to an ergative-absolutive pattern, as in Wambaya (see (5)). Comparing the distribution of case marking across these two classes, one is able to set up a system of three cases: ergative, nominative-absolutive, and accusative, whereby the nominative-absolutive is syncretic with the ergative for some words and with the accusative for others (see Goddard 1982).

(3) Yurok (Algic; California; Robins 1958: 20–1)

	'we'	cf. 'I'
NOM	*nekah*	*nek*
ACC	*nekah*	*nekac*

(4) Basque (Saltarelli et al. 1988: 211)

	'you.PL'	cf. 'we'
ABS	*zuek*	*gu*
ERG	*zuek*	*guk*

(5) Wambaya (West Barkly; Northern Territory, Australia; Nordlinger 1998: 80–4, 126)

	'grandmother'	'they'
ERG	*gugugayi*	*irriyani*
NOM–ABS	*guguga*	*irriyani*
ACC	*guguga*	*irra*

2.2 Syncretism involving the non-core cases. In the sample, syncretism involving a non-core case presupposes either syncretism involving a core case (§2.1) or the complete absence of core case distinctions (attested in Irish and Krongo), which suggests the following empirical universal: syncretism involving a non-core case implies the lack of core case distinction in some or all nominals.

Some instances involve syncretism of a non-core case with one of the core cases, typically those associated with transitive clauses, namely accusative (example (6)) or ergative (example (7)).

(6) Finnish (Fromm 1982: 89; foldout 1)

	'lock'	cf. 'I'
NOM	*lukko*	*minä*
ACC	*lukon*	*minut*
GEN	*lukon*	*minun*

(7) Burushaski (isolate; Pakistan or India; Klimov 1970: 42)

	'boy'	cf. 'woman'
ABS	*hiles*	*gus*
ERG	*hilese*	*guse*
GEN	*hilese*	*gusmo*

It is also possible for non-core cases to be syncretic with each other, involving either partial or complete collapse of non-core case distinctions. Partial collapse is well attested in Russian, where a number of different patterns are found, e.g. locative=dative (singular of *a*-stem nouns), locative=genitive (plural of adjectives and pronouns), genitive=locative=dative (singular of *i*-stem nouns), etc. Complete

collapse of non-core distinctions is illustrated by Ingush: attributive adjectives distinguish two forms, nominative and oblique, e.g. *joqqa* 'big.NOM' versus *joqqača* 'big.OBL', with the oblique corresponding to all seven of the remaining distinct case forms found in nouns (Nichols 1994: 99).

3 Geographical and genealogical tendencies

There are 74 languages of the sample which mark case inflectionally, with case syncretism found in 40 of them. It is most regularly found in Eurasia and Australia. The concentration in Eurasia is due to the presence of Indo-European, Uralic, and languages of the Caucasus (Nakh-Daghestanian and Kartvelian); these languages are also responsible for most of the examples involving syncretism among non-core cases (§2.2). Syncretism is especially well represented in Indo-European, where it is practically ubiquitous. The cluster of languages with case syncretism in Australia is largely, though not solely, due to the split ergativity pattern typical of Pama-Nyungan.

4 Theoretical implications

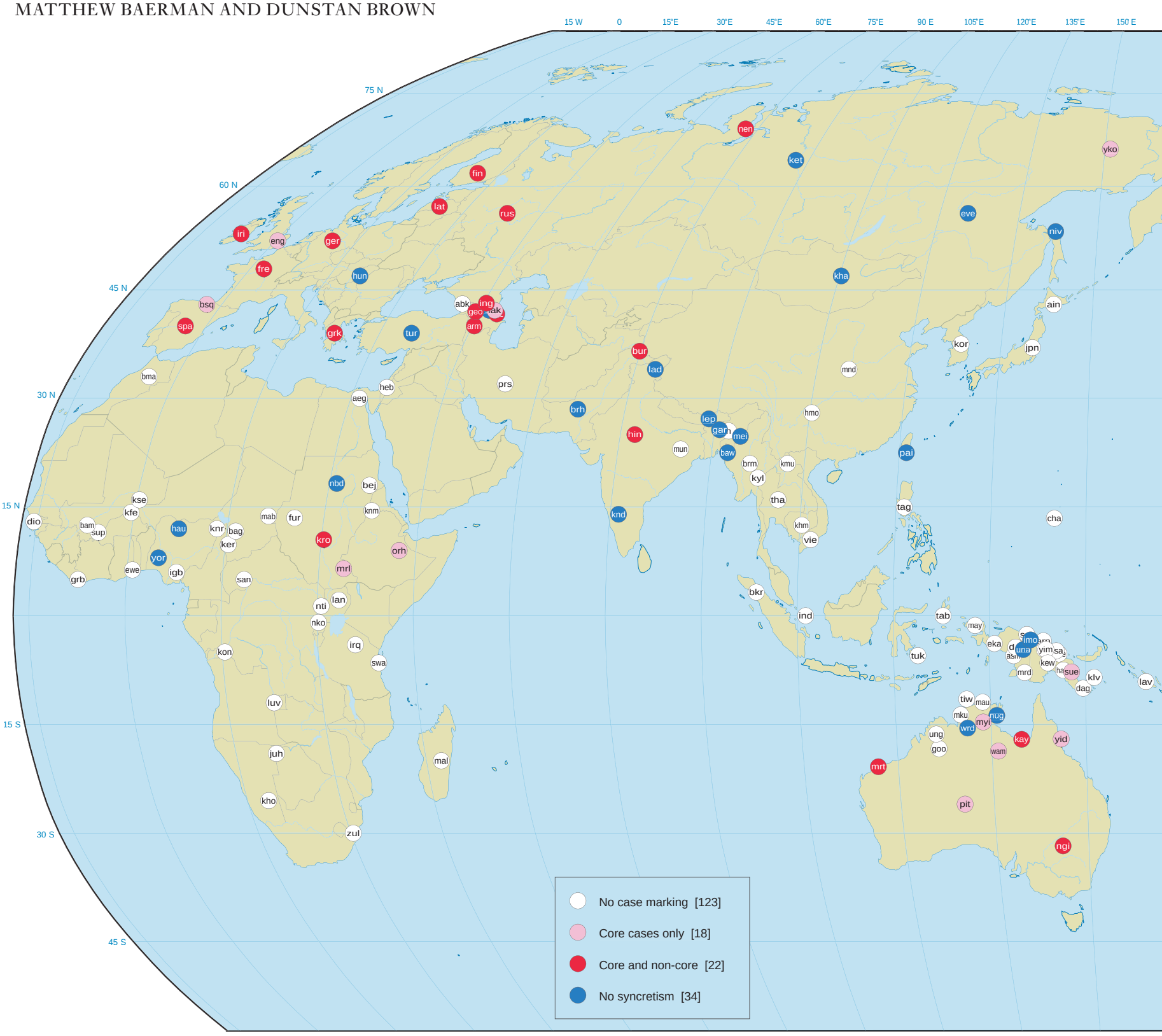
Linguists have long sought to use case syncretism as a window into the underlying structure of case semantics. If two cases can be expressed by a single form, so the reasoning goes, this must be because they share some element of meaning. A famous early example of this is Jakobson (1984 [1936]), where the multiple patterns of case syncretism found in Russian are used to set up a network of intersecting case features; this in turn owes a debt to earlier studies of Indo-European case systems, where case functions tended over time to collapse into an ever smaller number of forms. What the data from this chapter make clear is that only a subset of the phenomena characteristic of Indo-European have parallels elsewhere.

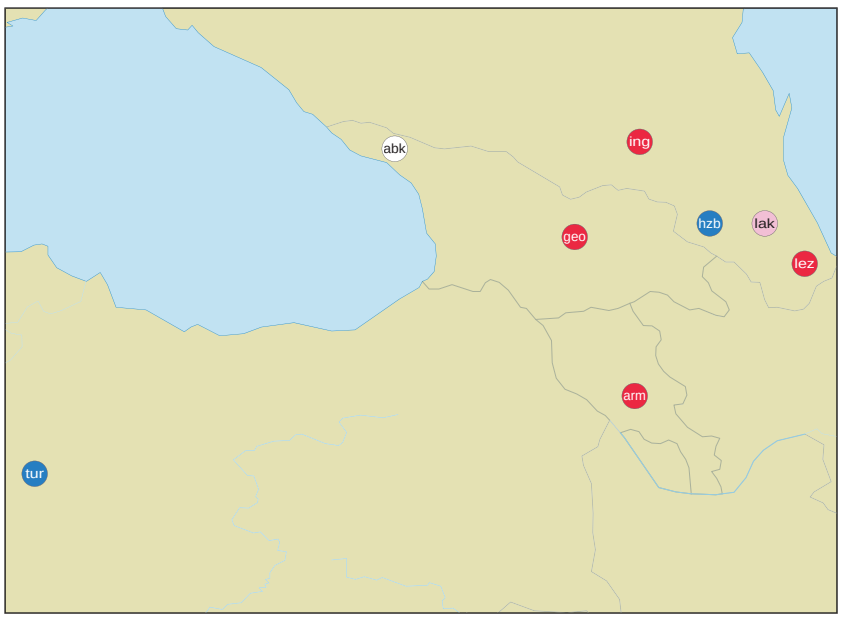
Outside of Indo-European, case syncretism almost exclusively involves a core case. The most common subtype, syncretism of the core cases with each other (§2.1), is often correlated with an animacy hierarchy, with personal pronouns at one end and inanimate nouns at the other. Higher animacy arguments are more likely to have a distinct accusative form, while lower animacy arguments are more likely to have a distinct ergative form. Syncretism of a core case with a non-core case (§2.2) seems to be an allied phenomenon: it looks as if it were a repair strategy to reverse the effects of core case syncretism, with the accusative or ergative "borrowing" the form of an oblique case in order to introduce a distinction between the core cases that would otherwise be absent from the paradigm. For example, in Eastern Armenian, where nominative and accusative are identical for inanimates, the dative case form is used to create an accusative for animate nouns (Minassian 1980: 90–1); while in Lak, where absolutive and ergative are identical for pronouns, it appears as if the genitive case form in *-ul* is used to create an ergative for nouns (Žirkov 1955: 64).

Syncretism involving solely the non-core cases (§2.2) shows few cross-linguistic regularities in the sample, and one suspects it is more a matter of language-specific morphological or phonological idiosyncrasies than general principles. Thus the elaborate networks of semantic relationships that have been sometimes constructed to account for syncretic patterns in Indo-European find little direct support from other language families, though the complete collapse of all non-core case distinctions, such as is found in Ingush, Georgian, and Russian, does appear at least to support the notion of core versus non-core as a morphosyntactic parameter.

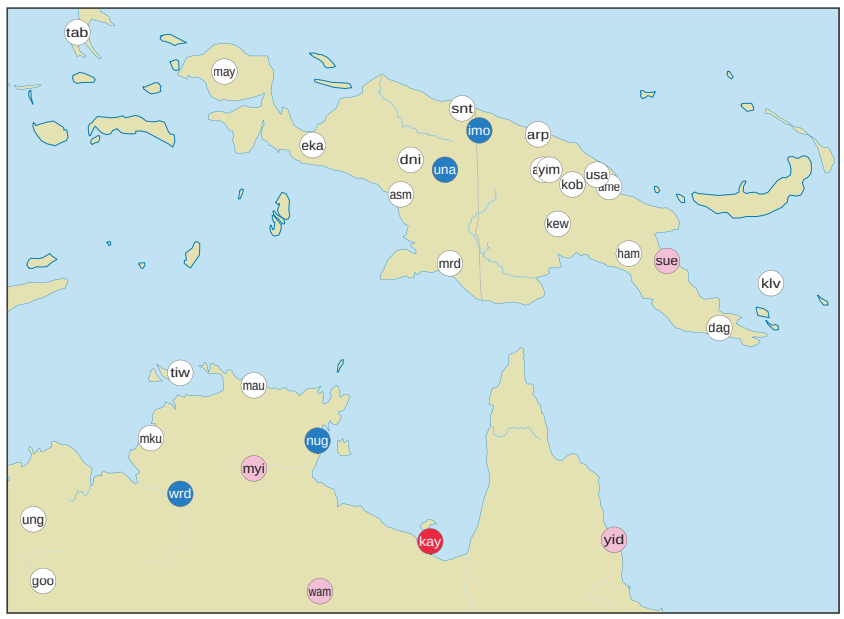
28 Case Syncretism

MATTHEW BAERMAN AND DUNSTAN BROWN

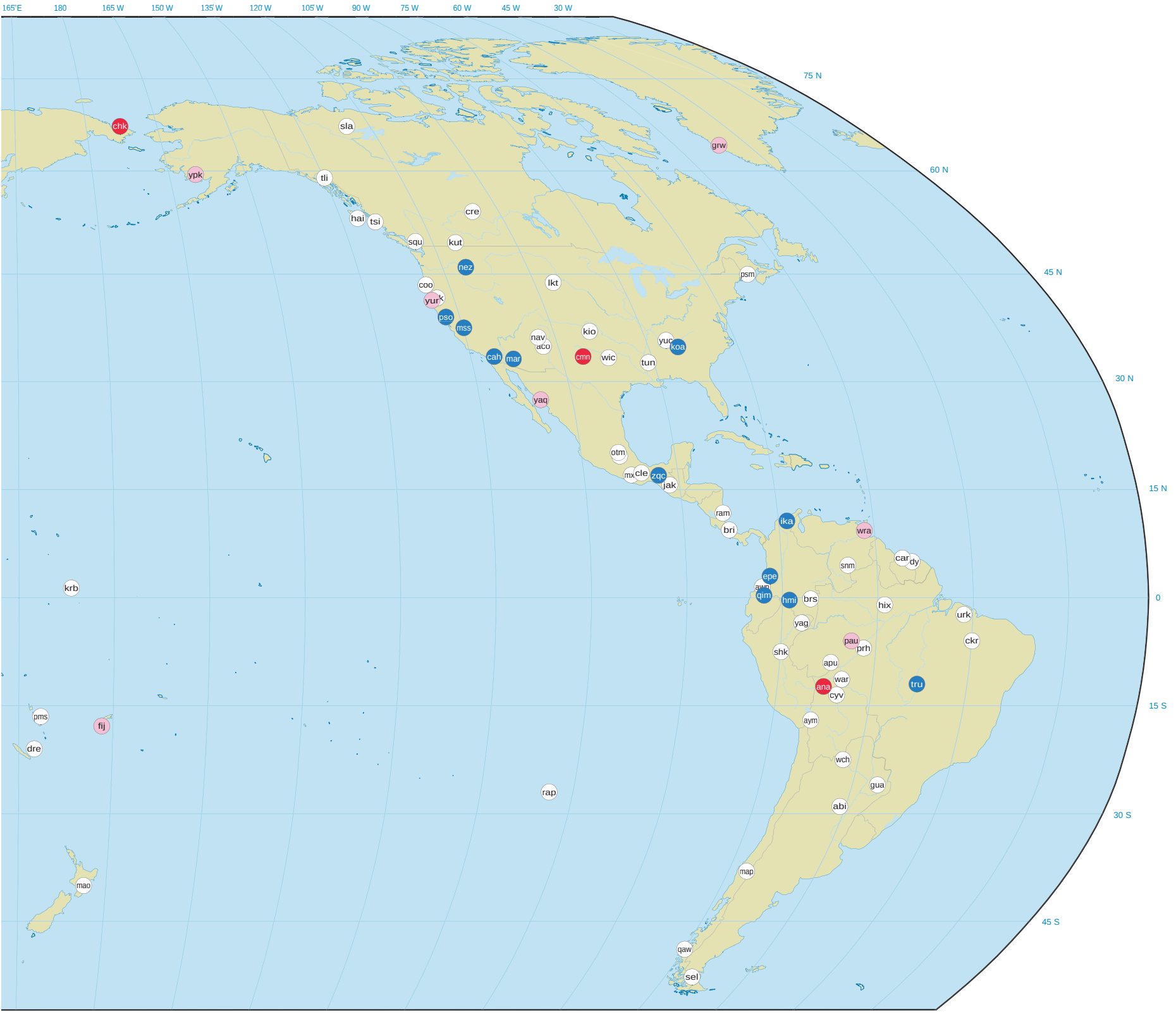


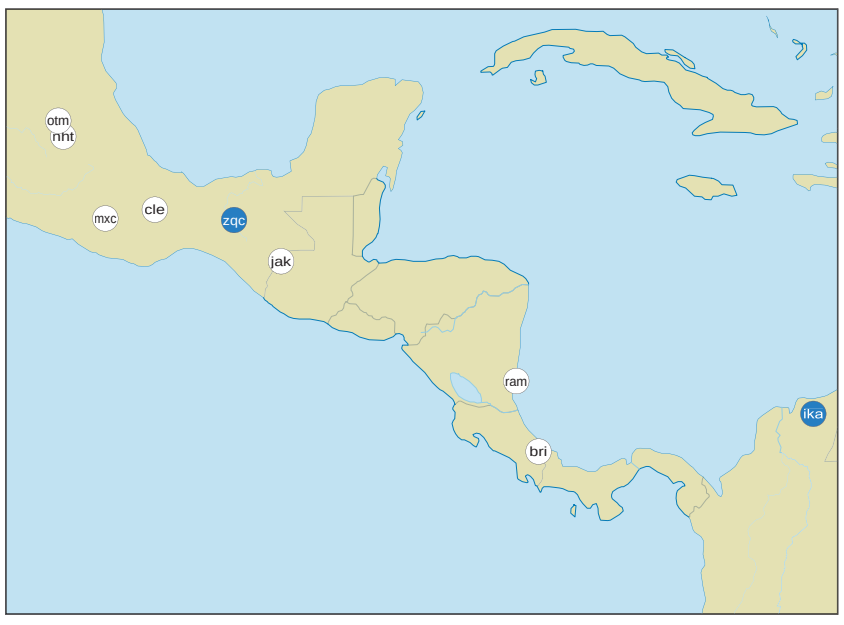


Caucasus and adjacent areas

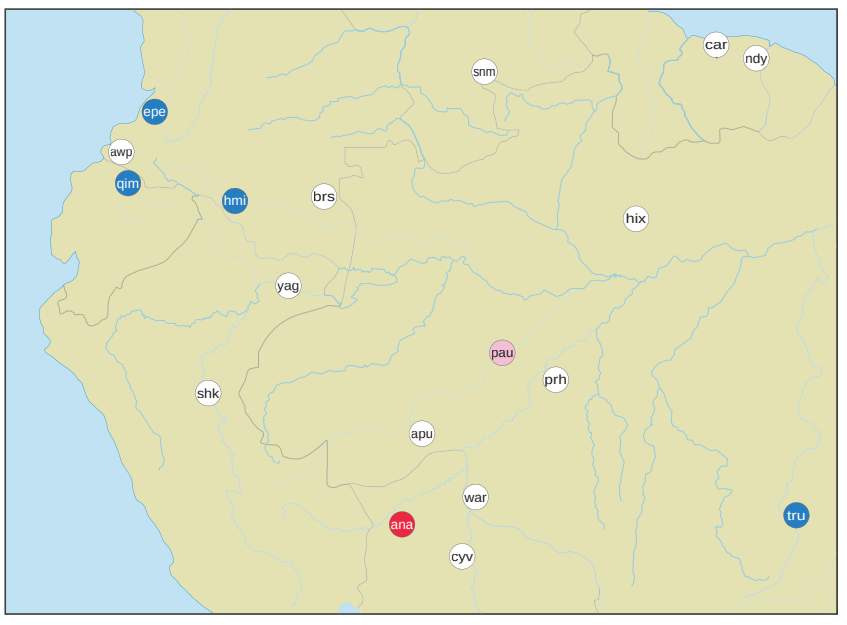


New Guinea and northern Australia





from southern Mexico to Columbia



northwestern South America

29 Syncretism in Verbal Person/Number Marking

MATTHEW BAERMAN AND DUNSTAN BROWN

1 Identifying verbal person/number syncretism

This map records instances of syncretism in the inflectional marking of subject person/number in verbs. For example, in German, the same form is regularly used for both first- and third-person subjects in the plural, e.g. *machen* '(we/they) make' or *machten* '(we/they) made'. The map makes a three-way distinction:

○	1. No subject person/number marking	57
●	2. Subject person/number is syncretic	60
●	3. Subject person/number is never syncretic	80
	total	197

In order to identify a form as syncretic one needs some evidence that there are multiple person/number values underlying a single form. This is straightforward when the syncretic pattern is somehow restricted in scope. For example, it may be found only in certain tense/aspect/mood paradigms, as in Aymara, where 1SG=2SG in the future perfect but not the present (see (1)); or it may be found only in certain inflection classes, as in Kunama, where 2PL=3PL in verbs of conjugation class I but not those of class IIb (see (2)); or it may be limited to a particular gender, as in Iraqw, where 2SG=3SG for feminine subjects but not for masculine ones (see (3)).

(1) Aymara (Bolivia and Peru; Deza Galindo 1992: 103–5)

'love'	future perfect	present
1SG	*munchiyäta*	*muntua*
2SG	*munchiyäta*	*muntahua*
3SG	*munchïna*	*munihua*

(2) Kunama (Nilo-Saharan; Eritrea; Bender 1996: 32–3)

aorist	class I 'sat'	class IIb 'had'
1PL	*gomake*	*maináke*
2PL	*goŋke*	*meináke*
3PL	*goŋke*	*oináke*

(3) Iraqw (Southern Cushitic; Tanzania; Mous 1993: 162)

'get up'	feminine	masculine
1SG	*tláw*	*tláw*
2SG	*tléer*	*tléer*
3SG	*tléer*	*tláy*

However, in some cases a given syncretic pattern is systematic throughout a language, so there is no direct evidence for the distinction. For example, in Kobon, 2DU and 3DU are always identical (see (4)). Nevertheless, there is indirect evidence, in as much as second and third person are distinct in the singular and plural; the same is true of the 1PL/3PL German form *machen* (see above), which corresponds to two forms in the singular, namely *mache* 'I make' and *macht* 's/he makes'.

(4) Kobon (Trans-New Guinea; Davies 1981: 166)

'went'	DU	SG	PL
1	*arlo*	*arnö*	*arno*
2	*arlö*	*arna*	*arbe*
3	*arlö*	*ara*	*arla*

In the remainder of this chapter, we will discuss a number of further generalizations that are not directly reflected on the map.

2 Common patterns of syncretism

Although it would appear that any combination of person values is a possible one, there are certain patterns which are more commonly attested than others. One relevant parameter is number. In the singular, by far the most common pattern is for second and third person to be identical, as in Iraqw (see (3)), while in the nonsingular, the most common patterns are for first or second person to be identical with third, as in German, Kobon (see (4)), or Kunama (see (2)). Another relevant parameter is the scope a given syncretism has within the language, i.e. whether the syncretism is limited in scope (as in (1)–(3)) or systematic throughout the language, as in Kobon (see (4)). Where syncretism is limited in scope, the singular is the most likely to be affected, accounting for more than half of the examples in the sample. On the other hand, where syncretism is systematic, it is rare for it to affect singular persons, accounting for only three examples in the sample (Hindi, Nivkh, and Dongolese Nubian), all of them involving 2SG=3SG.

The rarest type of syncretism is that where a single form combines different person values with conflicting values for another feature, typically number or gender. For example, in the "full form" paradigm in Diola-Fogny (used in certain syntactic contexts), the prefix *nu-* is used for both 2SG and 1PL (see (5)). In the prefix conjugation of the Semitic languages, 2SG masculine may be identical to the 3SG feminine, as in Egyptian Arabic *tiktib* 'write(s)' (Mitchell 1962: 73). Such examples are not numerous enough to reveal any striking cross-linguistic tendencies.

(5) Diola-Fogny (Northern Atlantic; Senegal; Sapir 1965: 90)

'broke'	SG	PL
1	*nitigɛr*	*nutigɛr*
2	*nutigɛr*	*jitigɛr*
3	*natigɛr*	*kutigɛr*

3 Geographical and genealogical tendencies

Person/number syncretism is widespread: of the 140 languages which mark person inflectionally (out of 197 total in the sample), roughly two-fifths (60 languages) display it. In most of these (54 languages) the syncretic pattern is restricted in scope, as shown in (1)–(3). The most syncretism-prone areas are Africa, New Guinea, western and northern Europe, and, to a lesser extent, South and Central America, and the Indian subcontinent. Least syncretism-prone are North America, Eurasia (with the exception of the above-mentioned parts of Europe), and the Pacific (with the exception of New Guinea).

In genealogical terms, person syncretism is widely found in nearly all of the sample languages from the Trans-New Guinea phylum, and is almost as widespread in the Nilo-Saharan, Niger-Congo, and Indo-European families. Some patterns can be said to be typical of certain families, at least as represented in the sample. Afro-Asiatic has 1SG or 2SG=3SG, restricted by gender. In Bantu, one of the gender/class markers used for third person is often identical to the marker for second or first person. In Carib languages, 1PL exclusive and third person are identical, opposed to a distinct first inclusive. In the Trans-New Guinea phylum, second and third person are identical in some or all nonsingular numbers.

4 Theoretical implications

A central question concerning syncretism is whether or not syncretic forms are indicative of some underlying semantic or morphological relationship. The data allow us to consider four possible interpretations.

First, one could assume that identity of form indicates identity of function, so that the syncretic form reflects a natural semantic or morphosyntactic class. For example, one could characterize the person values in terms of component features, so that the identity of first and second person in the singular in Aymara (see (1)) would be attributed to a shared feature such as [+discourse participant], while the identity of second and third person in the dual in Kobon (see (4)) could be attributed to a feature such as [–speaker].

Second, the syncretic form could be construed as unspecified for person. An obvious example inviting this interpretation is English, where the ending *-s* signals a 3SG subject in the present, while the bare stem is used for all other person/number values. Similar examples occur with first or second person: an overt 2SG marker contrasting with no marker for the other singular persons is found in the non-past in Ika (*n*^-; Frank 1990: 51), in the preterite in German (*-st*), and in various tense/aspect/mood paradigms in Spanish (*-s*); while an overt 1SG contrasting with no marker for other singular persons occurs in negative verbs in Karok (*ná-*; Macaulay 1992: 183).

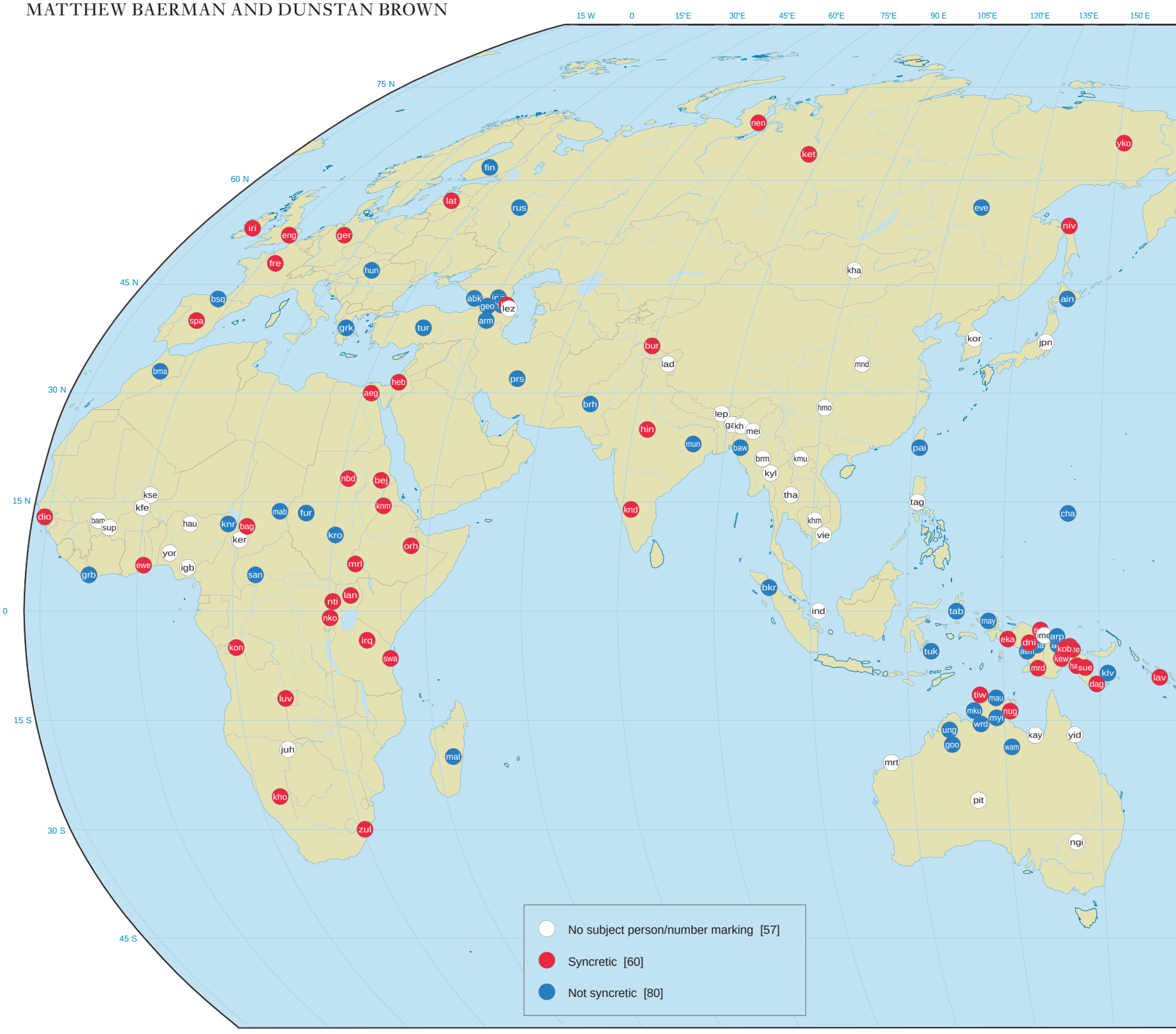
Third, one could say that the syncretic form actually "belongs" to one of its component values. Such an analysis is suggested by examples where a form which is nonsyncretic in one paradigm is used syncretically in another. For example, in Iraqw (see (3)), the syncretic 2SG / 3SG form used with feminines is identical to the distinct 2SG form used with masculines. There are other similar examples in the sample, which likewise involve second and third person, both in the singular (Igbo) and the plural (Carib, Lower Grand Valley Dani, and Suena).

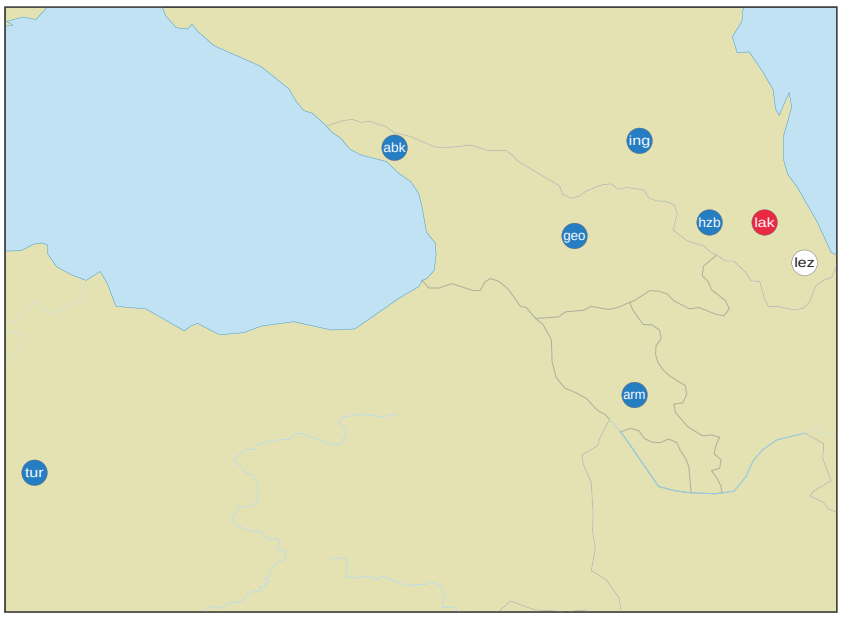
Fourth, syncretism could be the result of purely fortuitous homophony. Such an interpretation is most obvious where the syncretism is attributable to a morphophonological rule. For example, in Daga, a form for 1SG/3PL is opposed to a form for 3SG in the past tense of class A verbs by the alternation *-a-* ~ *-e-*; this is neutralized in class E verbs, which have nonalternating *-e-* throughout (Murane 1974: 52–3).

The data from the sample do not fully support any one single analysis. Each of the first two, by themselves, would be overly restrictive, given the range of patterns discussed in §2. The third possible analysis finds direct support only in a small number of cases. The fourth fails to take account of the existence of patterns which recur cross-linguistically. We suspect that all the analyses are valid in principle, being appropriate for different sorts of data.

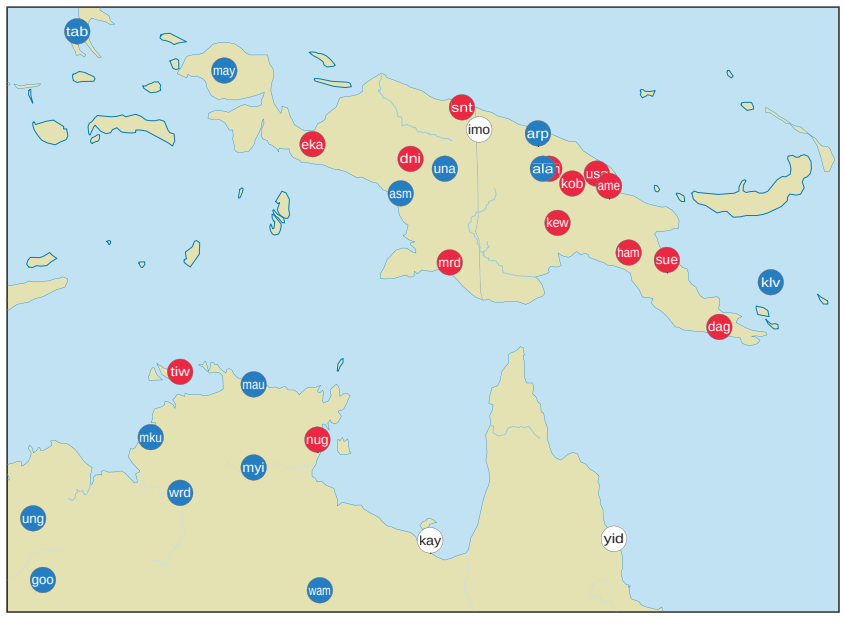
29 Syncretism in Verbal Person/Number Marking

MATTHEW BAERMAN AND DUNSTAN BROWN

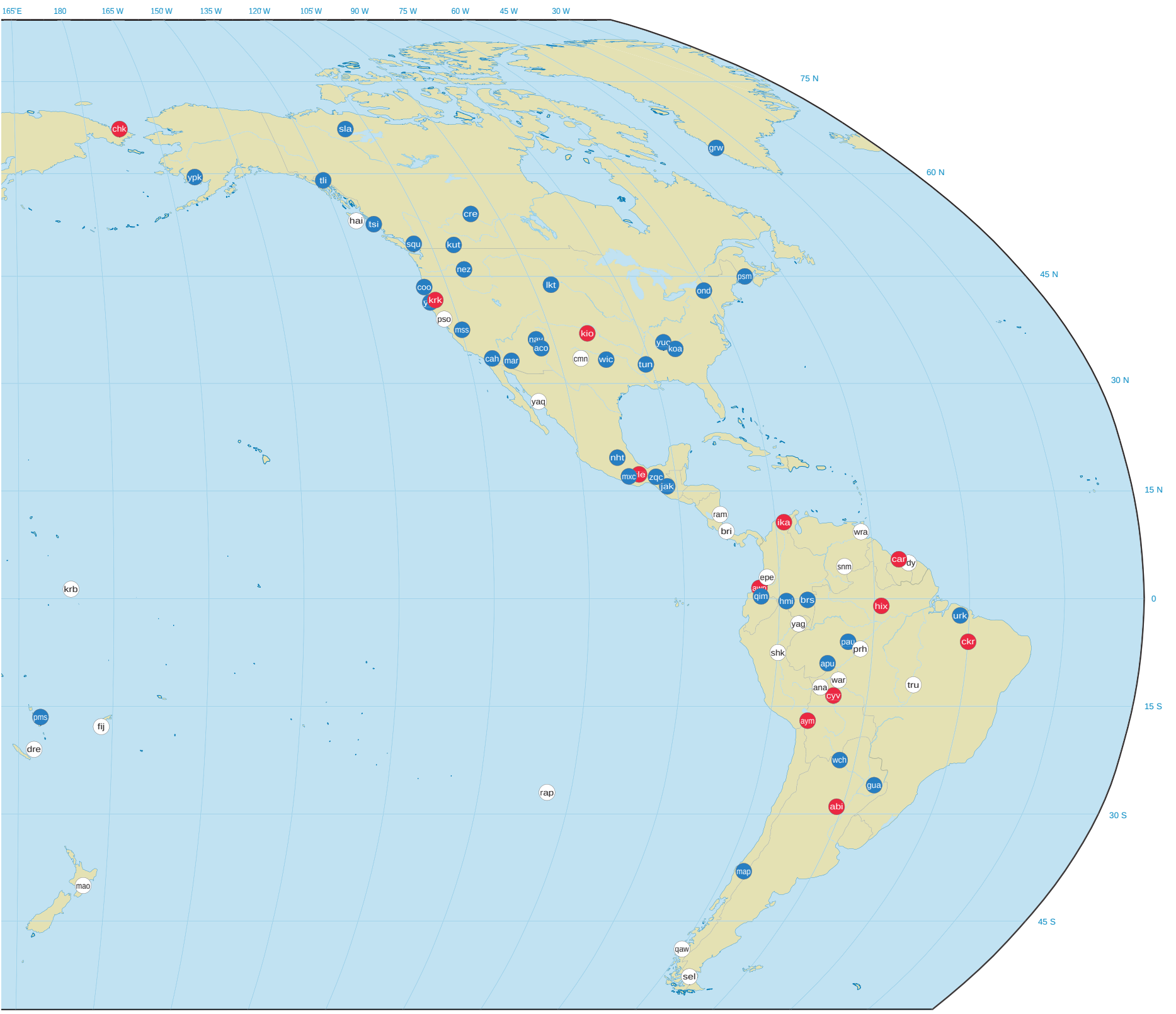




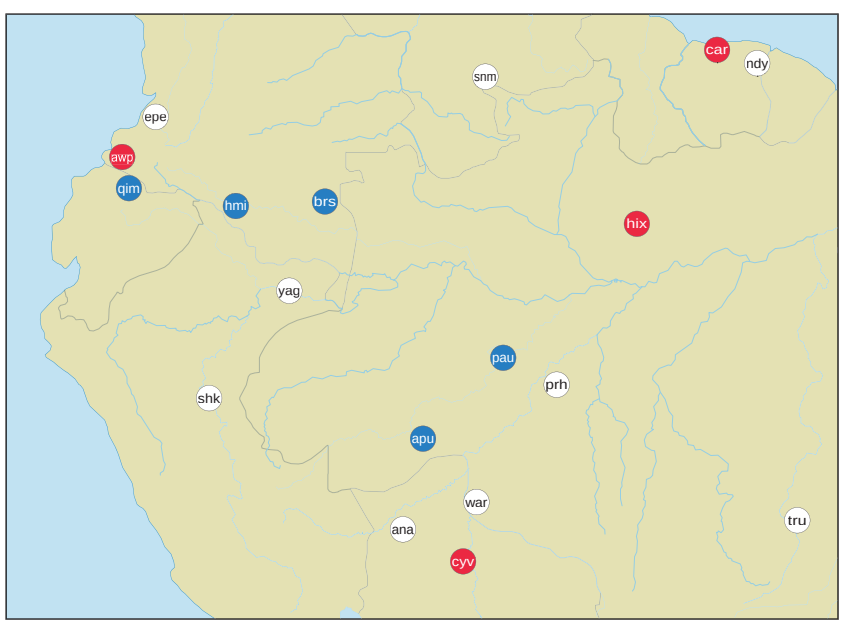
Caucasus and adjacent areas



New Guinea and northern Australia



from southern Mexico to Columbia



northwestern South America

30 Number of Genders

GREVILLE G. CORBETT

In some languages gender is evident in almost every phrase, while in other languages it is absent. This contrast raises interesting questions as to its function. Equally, in some languages it is relatively easy to determine the type of gender system and to establish how many genders the language has, while in others it takes careful analysis.

1 Defining the values

The defining characteristic of gender is **agreement**: a language has a gender system only if we find different agreements ultimately dependent on nouns of different types. In other words, there must be evidence for gender outside the nouns themselves. These Russian sentences illustrate the point:

(1) Russian

a. *Žurnal ležal na stole.*
magazine lay.M on table
'The magazine lay on the table.'

b. *Kniga ležal-a na stole.*
book lay-F on table
'The book lay on the table.'

c. *Pis'mo ležal-o na stole.*
letter lay-N on table
'The letter lay on the table.'

We find *ležal* 'lay' in (1a) as opposed *to ležal-a* or *ležal-o* because of the gender of the noun *žurnal* 'magazine'. The difference in the form of the verb in our three examples results just from the type of noun; there are no other differences in number, case, or syntactic structure. We are therefore dealing with a gender system. If in (1a), instead of the noun *žurnal* 'magazine' we had the noun *mal'čik* 'boy', *djadja* 'uncle', or some other noun denoting a male, the same form of the verb would be used. Hence nouns like *žurnal* 'magazine', as indeed nouns like *mal'čik* 'boy' and *djadja* 'uncle', are said to belong to the masculine gender. Nouns like *sestra* 'sister' would take the agreement as in (1b), and so nouns like *kniga* 'book' and *sestra* 'sister' are said to belong to the feminine gender. That leaves nouns like *pis'mo* 'letter' as the members of the neuter gender.

To avoid confusion note that the mere existence of nouns like *djadja* 'uncle' and *sestra* 'sister', denoting males and females, is not enough to constitute a gender system. There must be syntactic evidence, in agreement. Kanuri (Nilo-Saharan; Nigeria) does not have a gender system, but does have lexical contrasts such as *tádà* 'boy, son' versus *férò* 'girl, daughter' (Hutchison 1981: 11, 38, 45). This is a matter of lexical semantics, and not a gender system. Such lexical oppositions may be instantiated through derivational morphology (English shows examples like *poet* and *poetess*). Again this of itself does not give a gender system. The reason is that there can be numerous similar oppositions, concrete versus abstract for example, none of which would be counted as grounds for postulating a grammatical category in the language in question. Similarly, inflectional markers on the nouns themselves are insufficient to ground a gender system: in our Russian examples, the inflection *–a* might seem to indicate feminine gender, but in fact *djadja* 'uncle' and *papa* 'daddy' are masculine, which we prove by the agreements they take, irrespective of their form. We shall limit ourselves to true gender systems which can be demonstrated on the basis of agreement evidence.

Our examples have involved agreement of the verb, but there are various other targets which may agree in gender, such as adjectives, determiners, numerals, and even focus particles. Most scholars working on agreement include the control of anaphoric pronouns by their antecedent (*the girl . . . she*) as part of agreement. If this is accepted, as we do here, then languages in which free pronouns present the only evidence for gender will be counted as having a gender system. Of course, such languages with **pronominal gender systems** have a much less pervasive system than those like Russian. Including them, however, makes little difference to the overall picture, since they are rare (the best known example is English, which is typologically unusual in this respect; another is Defaka (Niger-Congo; Niger Delta, Nigeria; Jenewari 1983: 103–6)).

Since agreement is the defining characteristic, gender can be distinguished from other classification systems, such as **classifiers**. Classifiers are of various types, a common one being numeral classifiers analogous to *head* in the English expression *eighty head of cattle*. Distinguishing gender from classifiers is justified in Dixon (1982) and Corbett (1991: 136–7), and examples can be found in Aikhenvald (2000). It is possible, if unusual, for a language to have both gender and classifiers. One such language is Tariana (North Arawakan; Brazil; Aikhenvald 1994), which has a gender system and three subtypes of classifier. Coexistence of gender and classifiers is found in other languages of our sample: in Retuarã (Tucanoan; Columbia; Strom 1992: 10–11, 34–6, 45–7) and in Tidore (West Papuan; Halmahera, Indonesia; van Staden 2000: 77–81). Ngan'gityemerri (Daly; northern Australia) shows the development from generic classifiers into genders (Reid 1997). A further consequence of the definition is that differences in use of language which depend on the sex of the speaker (lexical choice, voice quality, and so on) are not treated here; an example is the difference between men's and women's pronunciation in Chukchi (Dunn 2000).

We should note that often there is no substantive difference between what are called "genders" and what are called "noun classes"; the different terms may be merely the products of different linguistic traditions. Thus we find systems with three genders, to which nouns are assigned by similar rules, in both Kannada (Dravidian; India) and Godoberi (Nakh-Daghestanian; eastern Caucasus). By tradition the first is said to have three genders, and the second three noun classes. We shall treat both as having gender.

We have established our criteria for deciding whether a language has a gender system. It is equally important to be clear on definitions when we ask how many genders particular languages have. Our approach starts from Zaliznjak (1964). Basically, two nouns are in the same gender provided that, however we change the environment (treating both the same), both will take the same agreements. Again traditions vary. The earlier Bantuist tradition treated nouns as being in different noun classes when singular and plural; we consider the total behaviour of a noun, including both its singular and its plural, with the result that a typical Bantu language may have 7–10 genders rather than around 20 noun classes. More generally, while in many languages there is no dispute as to the number of genders, there are a few where the question is far from straightforward. The analytical problem of determining the number of genders and the tests for deciding the gender of a given noun depend on separating out the classes into which nouns are divided (the controller genders) from the number of different genders marked on agreement targets (the target genders). Frequently the two match up, but in several languages they do not. A full treatment of the subject with extensive references can be found in Corbett (1991). Based on the analysis there, the analytical decisions made for this chapter are that the number of genders given on the map is the core system, the number of controller genders. (Hence neutral genders, locative genders, subgenders, overdifferentiated targets, inquorate genders, hybrid nouns, and those with double or multiple gender are all left out of account here; the interesting detail can be found in Corbett 1991: 145–88.)

The values for the number of genders are as follows:

1. None	144
2. Two	50
3. Three	26
4. Four	12
5. Five or more	24
total	256

As the figures show, the map is based on 256 languages, of which somewhat over half (144) have no gender system. A minimal gender system requires two genders, and such systems are common, with 50 examples in this sample. Three genders is around half as common (26 examples) and four genders, about half as common again (12). Larger systems, with five or more genders, are treated together, and represent a substantial minority (24 languages in the sample).

2 Geographical distribution

We look first at the distribution of gender languages versus non-gender languages. The two largest families differ in this regard. Niger-Congo has almost 1500 languages, a large proportion of which have gender systems (of interesting types, as we shall see in Chapter 31). This family is found in western, central, and southern Africa. The sample has 23 Niger-Congo languages, including Chichewa, Kisi, and Swahili. Just three in our sample have lost gender, following radical simplification of the morphological system; these are Ewe, Igbo, and Yoruba. The other very large family is Austronesian, with some 1250 languages spread across the Pacific, including Rapanui (the language of Easter Island), Indonesian, and Maori. Twenty-one Austronesian languages figure in the sample, none with a gender system, apart from the curious exception of Tagalog, which has partially borrowed the Spanish system (Schachter and Otanes 1972: 197–8). Elsewhere in Austronesian, gender has arisen in Teop (North-West Solomonic, Oceanic; Papua New Guinea; not in our sample) by three spatial demonstratives forming a system of gender-distinguishing articles (Mosel and Spriggs 2000).

While Niger-Congo is the major source, other families of Africa are gender "hotbeds" in Nichols's terms (1992: 130–2). Khoisan languages in the south (represented by Ju|'hoan in our sample) have interesting gender systems (see Güldemann 2000), and Afro-Asiatic in the north contributes many gender languages (including Amharic, Modern Standard Arabic, Hausa, Hebrew, and Qafar). Nilo-Saharan is mixed, but none of the five languages in the sample shows gender. To the north, Europe is a predominantly gender area, with Indo-European languages such as French, German, and Russian. In the Caucasus, the Nakh-Daghestanian family is a stronghold of gender. Indo-European extends to the South Asian subcontinent, with gender languages such as Hindi and Marathi. In southern India, Dravidian languages typically show gender (examples in our sample are Kannada, Tamil, and Kolami). Elsewhere gender is less strong. Austro-Asiatic presents a mixed picture, with gender in Khasi and Nicobarese, but not in Khmer or Vietnamese. In New Guinea, several families show gender, and of widely differing types. In Australia there are several gender languages, such as Maung and Bininj Gun-Wok, mainly clustered in the north. In the Americas, gender languages form a minority. In North America, there are Algonquian languages like Plains Cree and Eastern Ojibwa (discussed in Chapter 31). There are a few gender languages in Central America, such as Chalcatongo Mixtec and Lealao Chinantec (both Oto-Manguean; Mexico). In South America the picture is mixed, with several of the families of Amazonia including gender languages. From the opposite perspective, the main areas *without* gender are the Pacific, most of Asia (notably the Sino-Tibetan family), including Siberia (notably the families grouped under Uralic and Altaic), together with several families of North America, especially those to the west.

Our discussion has been in terms of languages; if we think of numbers of speakers, then the major gender family would be Indo-European, which includes English, French, German, Hindi, Portuguese, Russian, Spanish, each with many millions of speakers. On the other, nongender side would be Sino-Tibetan, in which Mandarin and Cantonese alone account for a substantial proportion of the world's population. Again, then, there is a substantial split. Many speak gender languages and many do not.

Given a gender system, the most common number of genders is two. Such languages are found in most gender areas. Afro-Asiatic systems regularly have two genders, though there are complex interactions with number (see Corbett and Hayward 1987). In Indo-European many languages retain three genders (like Icelandic and German), while many others have reduced to two (like French and Spanish); a minority have lost gender altogether (e.g. Eastern Armenian). Four-gender systems are particularly prevalent in Nakh-Daghestanian languages (our sample includes Archi, Lak, and Tsez), though they occur elsewhere too, as in the isolate Burushaski. For larger systems the major source is Niger-Congo, where systems in excess of five genders are common. Nigerian Fula is exceptional, having around twenty genders, depending on the dialect (Arnott 1967; 1970: 67–75; Koval' 1979; Breedveld 1995: 295–460). Other large systems can be found in Papua New Guinea, where Arapesh has thirteen genders (Fortune 1942; Aronoff 1994: 89–114; Fraser and Corbett 1997). In northern Australia, too, Ngan'gityemerri arguably has fifteen genders (Reid 1997).

3 Theoretical issues

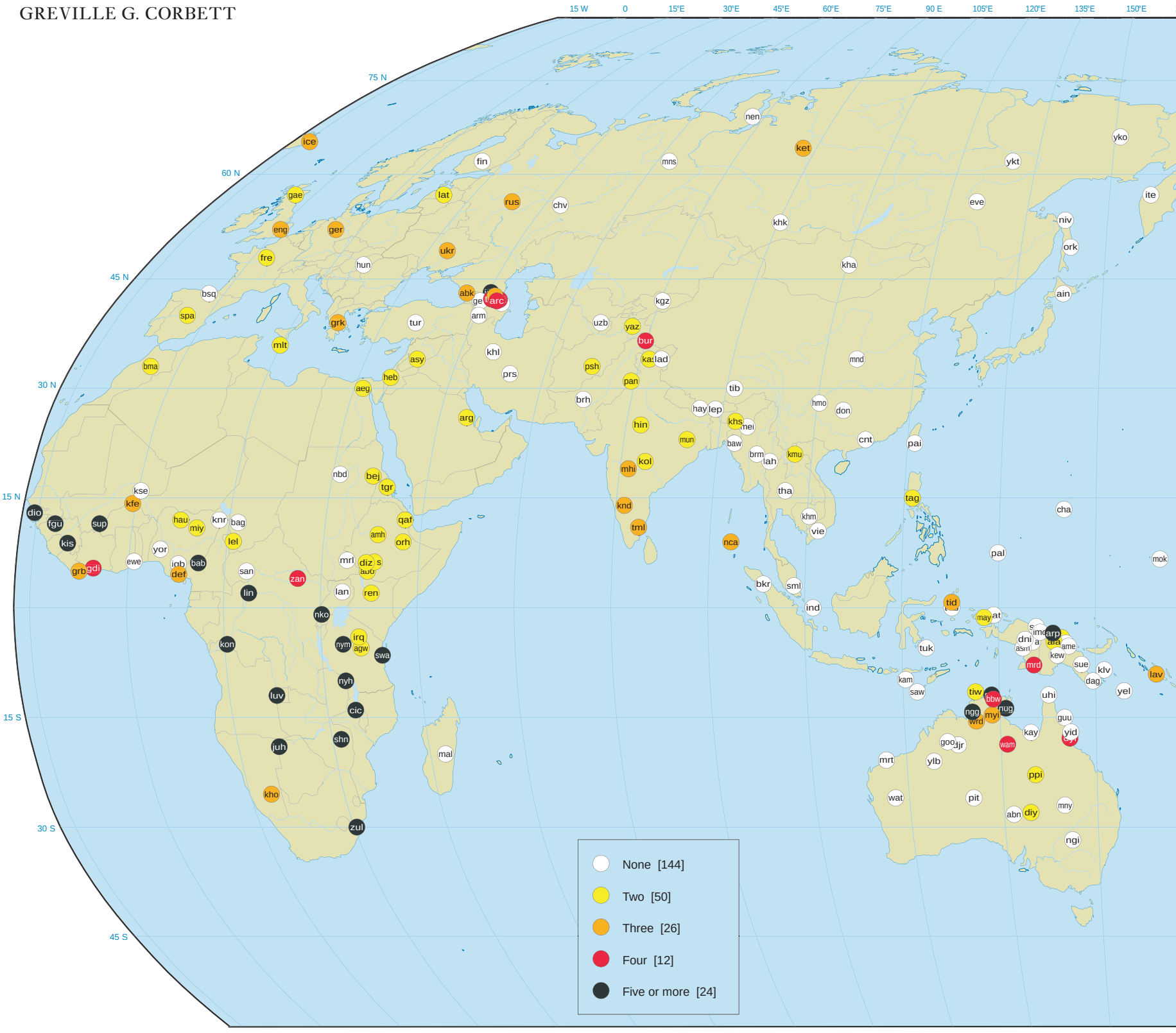
One substantial issue, noted earlier, is the function of gender systems, given that they pervade some languages, are of some importance in others, and are missing in yet others. This and related issues are discussed in Corbett (1991). Besides that source, the questions of the way in which gender systems develop, and the way in which they are represented, are also treated briefly in Chapters 31 and 32.

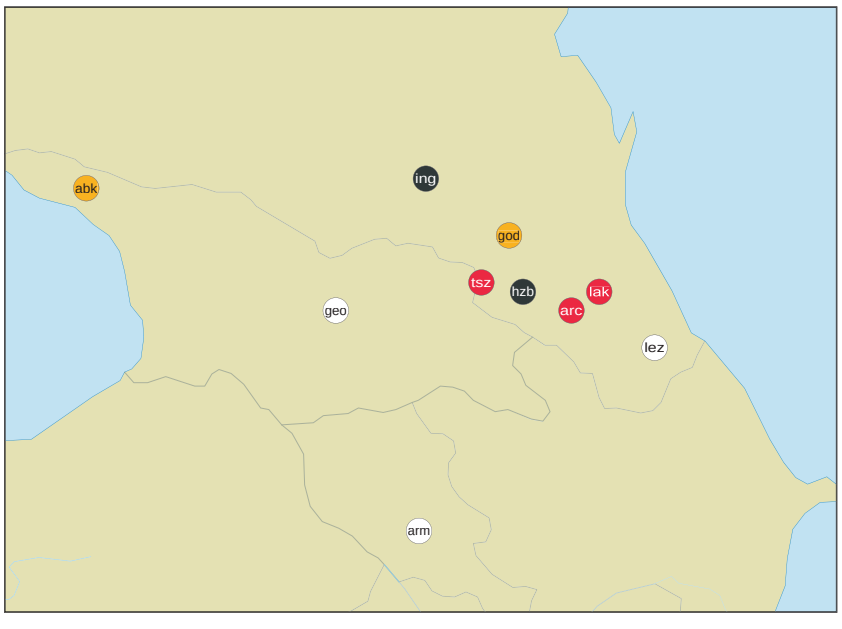
Acknowledgement

This research was funded in part by the ESRC (UK) under grants R000271235 and RES 051270122. This support is gratefully acknowledged.

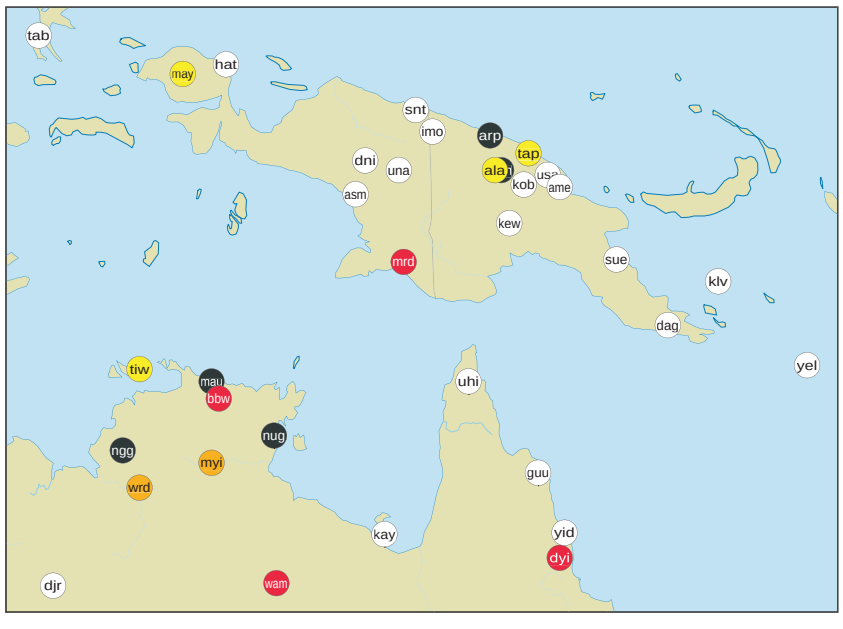
30 Number of Genders

GREVILLE G. CORBETT

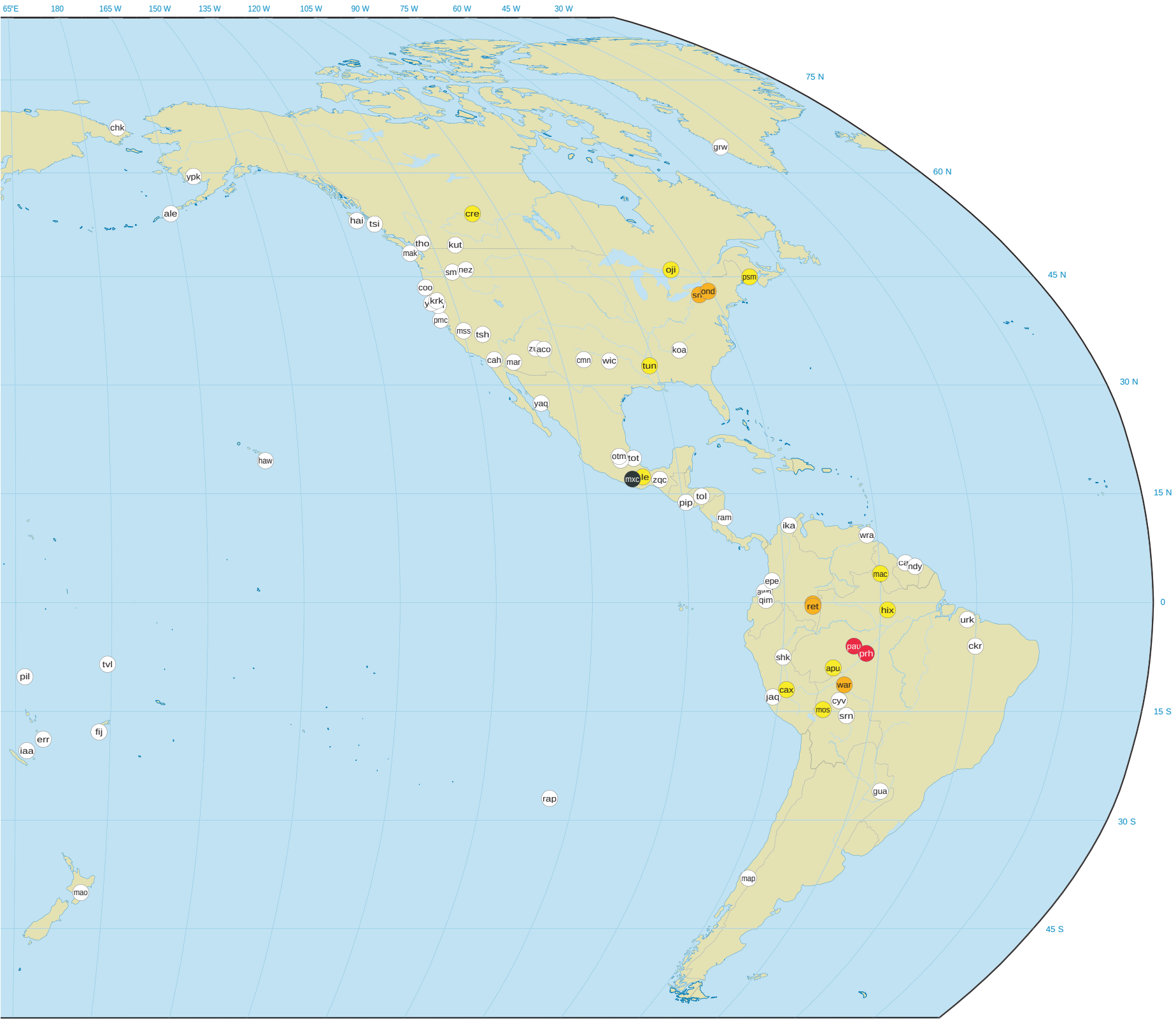


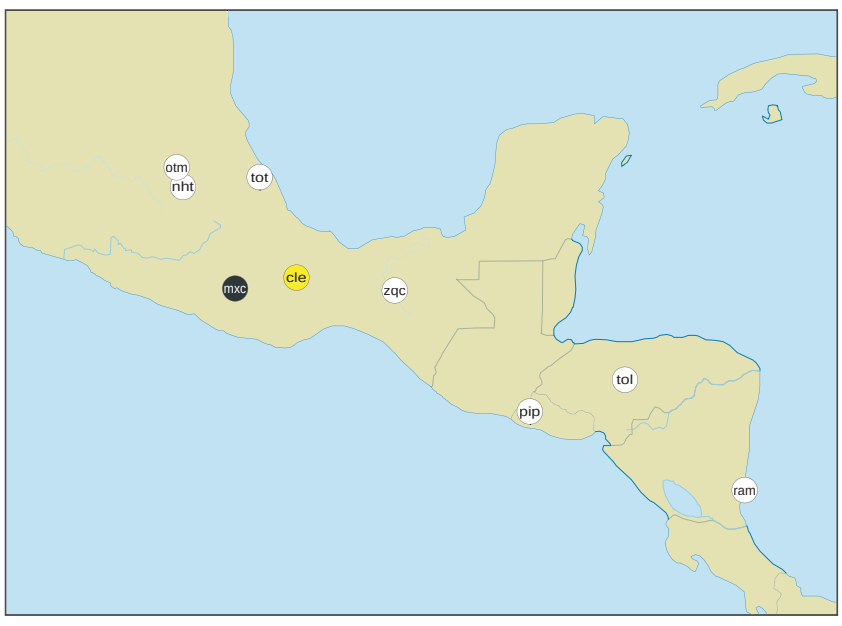


Caucasus

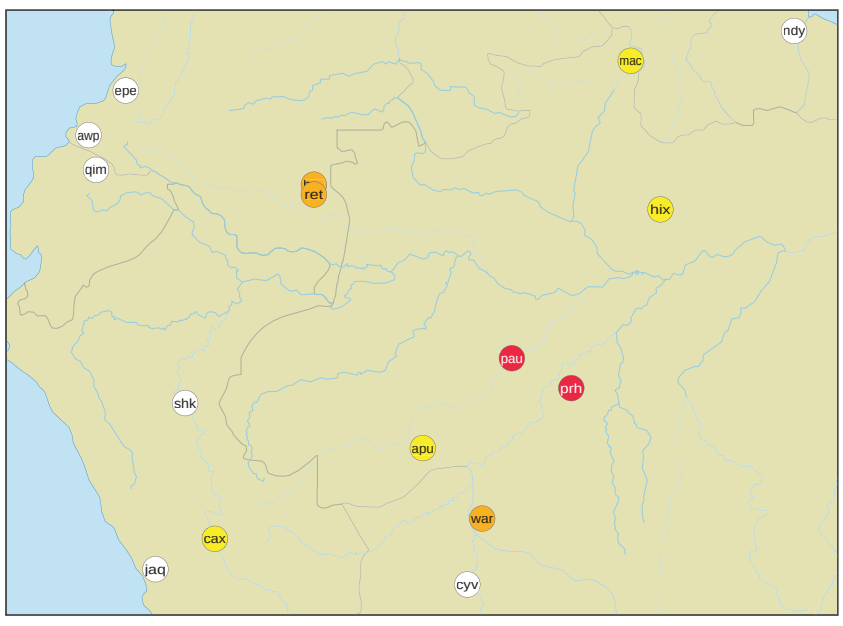


New Guinea and northern Australia





southern Mexico and Central America



northwestern South America

31 Sex-based and Non-sex-based Gender Systems

GREVILLE G. CORBETT

Linguistic gender systems are frequently linked to biological sex. This is not the only possibility; alternatives occur, particularly in some of the larger gender systems.

1 Defining the values

We saw in Chapter 30 how, in many languages, nouns may be divided into groups according to the agreements they take, even when we control for other factors such as number and case. We should then ask whether these groups are arbitrary. The answer is that there is always a semantic "core" to the system. That is, there is an overlap between the nouns which take a particular set of agreements and some semantic feature. (This overlap may be greater or smaller, as we shall see in Chapter 32.) In the familiar systems such as French and German, and indeed in the majority, the link is to biological sex. This was illustrated from Russian, in Chapter 30, and can be seen again in Bininj Gun-Wok (Gunwinygic; northern Australia; Evans et al. 2002: table 1). Here we see agreement of the adjective in gender in the Kunwinjku dialect of the language (older speakers):

(1) Bininj Gun-Wok

a. *bininj* *na-mak*
man I-good
'good man'

b. *daluk* *ngal-mak*
woman II-good
'good woman'

c. *kamarn* *man-mak*
cheeky.yam III-good
'good cheeky yam'

d. *kukku* *kun-mak*
water IV-good
'good water'

Nouns denoting males, like *bininj* 'man', are found in gender 'I', and take the 'I' agreement marker *na-*. Male sex forms the semantic core of this gender, but there are other nouns found in it too; for example, it is the default gender for nonhuman animates. Nouns denoting females, like *daluk* 'woman', are found in gender 'II', and take the 'II' agreement marker *ngal-*. Again, though female sex forms the semantic core of this gender, there are other nouns in it too, mainly denoting lower animates.

Such systems present considerable variety, as we shall see. However, not all gender systems are sex-based. Consider these examples from Maasina Fulfulde (Atlantic; Breedveld 1995: 295):

(2) Maasina Fulfulde

a. *nɛɗɗɔ* *ʔɔ*
person DEF.'O
'the person'

b. *ɲiiwa* *ba*
elephant DEF.BA
'the elephant'

c. *laana* *ka*
boat DEF.KA
'the boat'

d. *lekki* *ki*
tree DEF.KI
'the tree'

Fulfulde/Fula has around twenty genders (dialects vary: Guinean Fula, the dialect recorded on the map, is particularly rich); the singular forms of some genders are illustrated here. They are usually labelled by capitals, according to the form of the article they take (e.g. 'O or BA); adjectives, demonstratives, numerals, and pronouns also show similar agreement. We can see that the form of the article changes to mark agreement with the noun, as in French or German. The noun has a class suffix, which may be similar in form to the article, or it may be in conflict with it (Breedveld 1995: 296), in which case, of course, it is the agreement which tells us the gender. The gender 'O/'BA (that is, nouns taking 'O in the singular and 'BA in the plural) includes *dɛbbɔ* 'woman', and indeed a whole range of nouns denoting humans. Thus it has a clear semantic core; however, sex is not a part of it: nouns denoting human males and human females are found in the same gender. Other genders overlap with semantic categories to a greater or lesser degree; thus together with 'elephant' we find other large animals, wild and domesticated, and the word for 'field'. Sex plays a minor role in the Fulfulde NDI gender, where according to Breedveld (1995: 329–30) there are nouns denoting signs of wealth, seeds, slave names, and domestic male animals; note, however, that not all domestic male animals are in this gender (for example, 'male donkey' belongs with 'elephant' (1995: 411)).

There are various types of gender systems where biological sex is not the semantic core. These are all based on some notion of animacy. When required there is no problem about expressing distinctions based on sex in such languages, by the use of separate lexical items (whether based on different roots or derivationally linked) or by qualifying items with the meaning 'male/female'. The use of the term *gender* is still fully appropriate for systems based on animacy, because structurally the systems are fully comparable. Note too that *gender* derives etymologically from Latin *genus*, via Old French *gendre*, and originally meant 'kind' or 'sort'.

The values for the map, which is based on the same sample of 256 languages as Map 30, are as follows:

○	1. No gender system	144
●	2. Sex-based gender system	84
●	3. Non-sex-based gender system	28
	total	256

2 Variety in sex-based systems

Though a major division is between sex-based and non-sex-based systems, we should not neglect the variety within these groupings.

First there is variety in how closely the grammatical gender system relates to the corresponding semantic category. At one end of the scale, in sex-based systems the genders may match the semantic category almost completely. Examples can be found in Dravidian languages, where for instance in Tamil, it is almost correct to say that nouns denoting male humans are masculine, and masculine nouns denote male humans. There is more to be said, since this gender also includes male deities, but the statement captures the essence of the system. Compare this with the very different system of a typical Indo-European language like French or Russian, where it is also correct to say that nouns denoting males are typically masculine, but where the masculine nouns include a large proportion of nouns which do not denote males. We may also find exceptions where nouns denoting humans of one sex are found in the 'wrong' gender. Typically, however, these nouns are not fully in that gender but take agreements of more than one type and are "hybrid nouns" (Corbett 1991: 176–81).

Second, these languages may set the threshold for "sex differentiability" at different points. Sex distinctions extend to insects and plants, but no language has been reported as including reference to their biological sex within a grammatical system. Humans are most interested in the sex of other humans, and the threshold may well be

set here. This is the case for Tamil, where both 'bull' and 'cow' are in the neuter gender. Similarly in Tidore (West Papuan; North Moluccas, Indonesia), a rooster is treated as grammatically neuter, as is a pregnant goat (van Staden 2000: 77–8 and personal communication). Other languages set the threshold lower. In Russian, sex-differentiability extends to creatures whose sex matters to humans (that is, primarily animals which humans breed), or where the difference is striking (as with lions); nouns which fall below the threshold may be in any gender (thus 'shark' is feminine and 'dolphin' is masculine).

There are some curious effects of status, usually within the feminine gender. Lak (Daghestanian; central Daghestan highlands) has four genders, in broad outline: male rationals (I), female rationals (II), other animates (III), though this has other members too, including many inanimates, and a residue gender (IV), which also includes a few animates. There was an important exception, namely *duš* 'girl, daughter', a member of gender III instead of the expected gender II. Gender III agreements became a sign of politeness when addressing young women (Xajdakov 1963: 49–50), particularly those earning their own living, and nouns denoting them have been transferred to gender III. This usage has extended so that now gender III agreement forms are appropriate for any woman outside the immediate family. Within the family, older women such as *ninu* 'mother' and *amu* 'grandmother' are addressed and referred to using gender II forms. Thus gender II is semantically restricted and is left with extremely few nouns in it. Something comparable has happened in Konkani (Indo-European; west coast of India; Miranda 1975: 208–13), where the word for 'girl' was neuter. Where human referents are concerned, the neuter has become the gender for young females (or those relatively younger from the speaker's standpoint), while the feminine is for old, or relatively older, females. A similar change in the core meaning of genders has occurred in some southern Polish dialects (Zaręba 1984–85). In several of these dialects, nouns denoting girls and unmarried women (irrespective of age), and including hypocoristics, are of neuter gender. Neuter agreements are employed when unmarried women are addressed, and they use them for self-reference. In a smaller area, to the south-west of Kraków, instead of the neuter the masculine is used. In both types of dialect, the feminine is used for married women. The change from neuter or masculine to feminine for a particular woman occurs immediately after the church wedding ceremony; the communities involved are small, and so there is no difficulty about knowing who is married and who is not (A. Zaręba, p.c.). The meaning of the feminine has changed in both dialect types, being restricted now to denote married women. (Feminine nouns which are not semantically motivated also remain feminine.) For further details on all these, and suggestions as to how they have arisen, see Corbett (1991: 24–6, 99–101).

3 Variety in non-sex-based systems

These systems are all based on some type of animacy. The first source of variation is the threshold for differentiability. One possible distinction is human versus nonhuman, as we saw in Fulfulde. This is found more widely in Niger-Congo; a clear account of these systems can be found in Welmers (1973: 159–83). Several Bantu languages have lowered the threshold, so that animals go together with humans to give an animate gender (this change in languages of the coastal area of Kenya and northern Tanzania is documented in Wald 1975). The other main area for non-sex-based gender systems is the Algonquian family of Canada and the northern United States. These two-gender systems distinguish animate and inanimate. As with sex-based systems, a source of variation is the type of "leaks" that occur into semantically motivated genders. In Algonquian there are various examples of apparent nonanimates which are treated as animates. For example, in Eastern Ojibwa, nouns denoting persons, animals, spirits, and trees are animate: *enini* 'man', *enim* 'dog', *menito*: 'manitou', *mettikumi:šš* 'oak' (Bloomfield 1957: 31–2). Others are inanimate. But several nouns are unexpectedly animate, including: *enank* 'star', *meskomin* 'raspberry', *ekkikk* 'kettle'. Various approaches to this unexpected deviancy in animacy could be taken. One might simply treat them as lexical exceptions; they are few in number, and such small groups of exceptions are often found in gender systems. Others have suggested that animacy is only a part of the explanation, and that the system must be seen in terms of a different world view, in which "power" is the dominant element, and where nouns treated as animate have at some point been viewed as denoting powerful entities. A proponent of this view is Black-Rogers (1982); see Corbett (1991: 20–4) for further data and sources.

A second source of variation is the nature of the semantic basis for the other genders. In larger systems, as in Niger-Congo, the nonhuman or nonanimate genders (of which there may be several) may in turn be semantically justified to a greater or lesser extent. This is an ongoing debate, in that some see the apparently mixed collection of nouns in certain genders as reflecting some conceptual cohesion, while others treat them as relics of earlier systems, now linked by formal rather than semantic similarity.

4 Geographical distribution

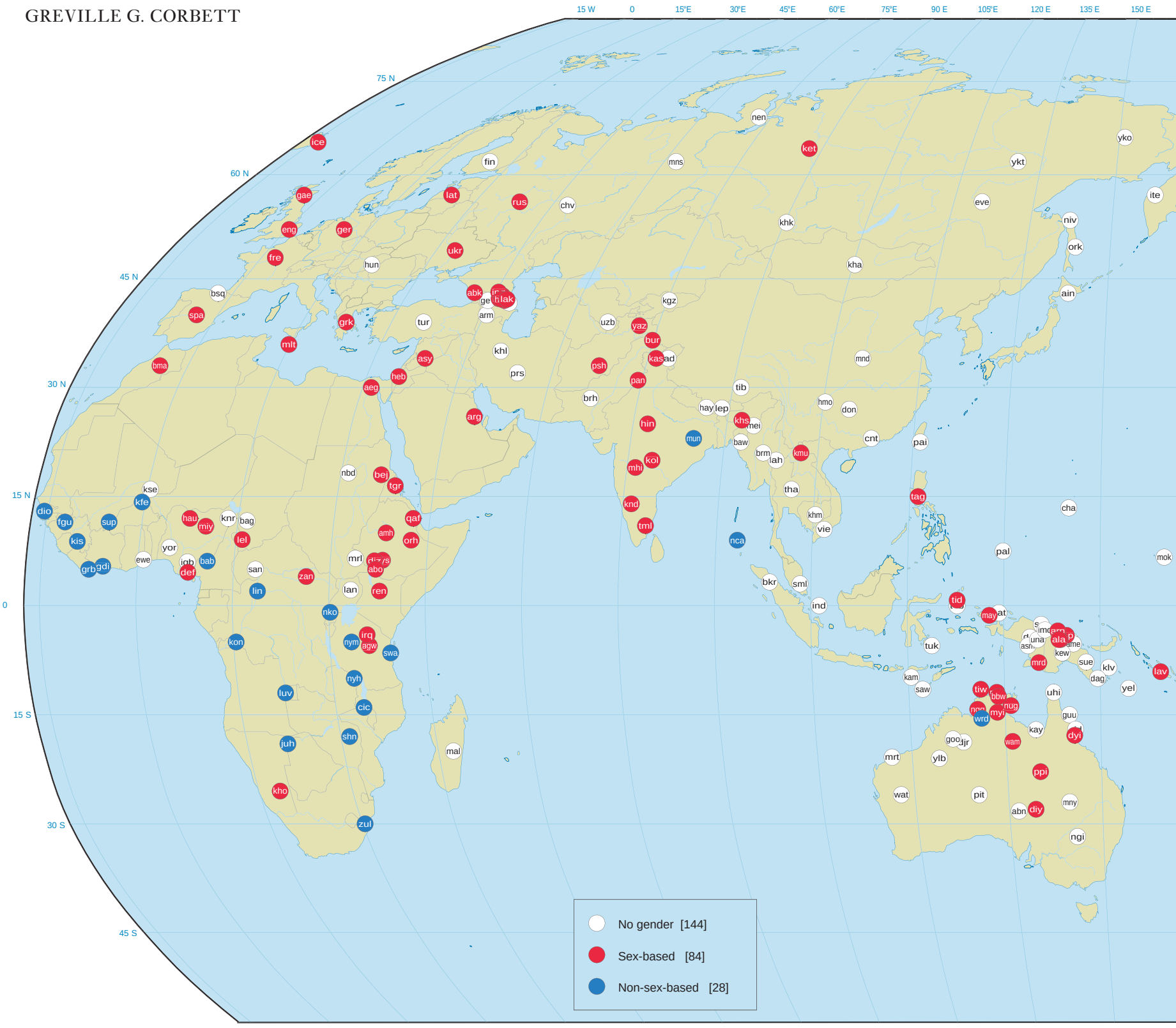
The picture is relatively clear. Sex-based systems are found in almost all areas where there is gender. Of the 112 languages with gender in the sample, three-quarters (84) have sex-based systems. The main non-sex-based area is covered by the extensive Niger-Congo family in western, central, and southern Africa, which contributes 17 of the 28 non-sex-based systems. Most of these have five or more genders, but Grebo and Koromfe have three. The other substantial non-sex-based area is that of the Algonquian family of North America, reflected in our sample by Plains Cree, Eastern Ojibwa, and Passamaquoddy-Maliseet. Elsewhere there is Ju|'hoan, representing Khoisan languages from southern Africa. In Austro-Asiatic, the languages in our sample are Mundari and Nicobarese. In Australia there is Wardaman. Two Carib languages (Hixkaryana and Macushi) are both of this type, as is Lealao Chinantec (Oto-Manguean; Mexico). The wide scatter of these languages shows that animacy is a viable basis for gender systems. Nevertheless, it is overshadowed by sex-based-systems.

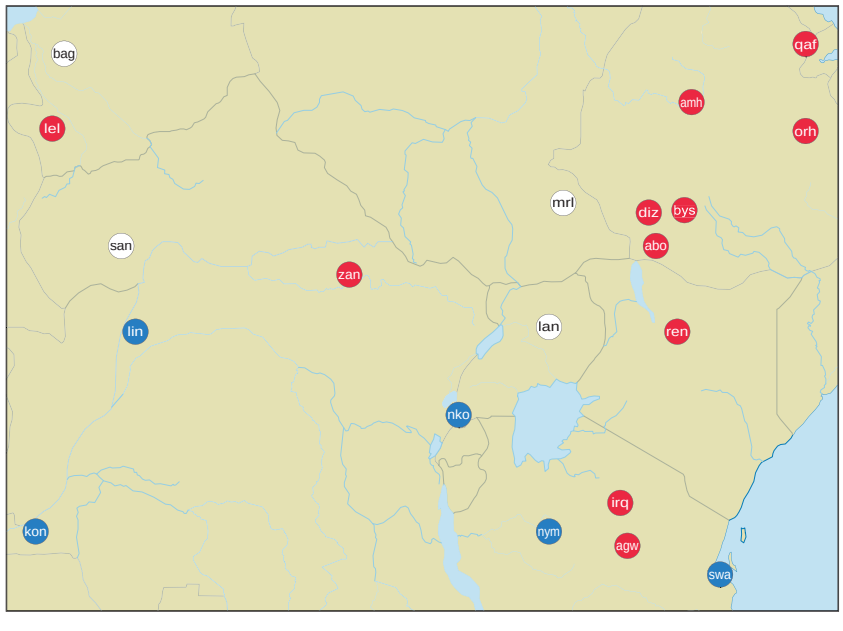
5 Theoretical implications

The existence of systems which are structurally similar, but which have different semantic content, offers exciting perspectives for psycholinguistic research. The longevity of the non-sex-based systems within Niger-Congo has implications for historical linguistics, since it shows that a minority system can maintain itself and develop in a variety of ways without adopting the majority scheme. Finally, for those interested in language and gender in the sociological sense, these direct reflections of biological sex in many languages, with the "control group" of languages with similar gender structures but without the sex component, should provide a valuable source of data.

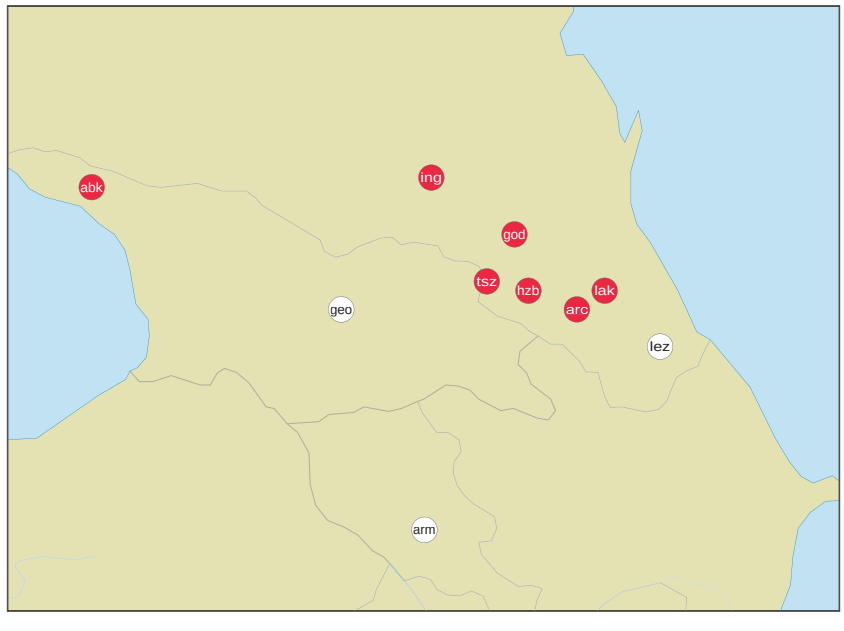
31 Sex-based and Non-sex-based Gender Systems

GREVILLE G. CORBETT

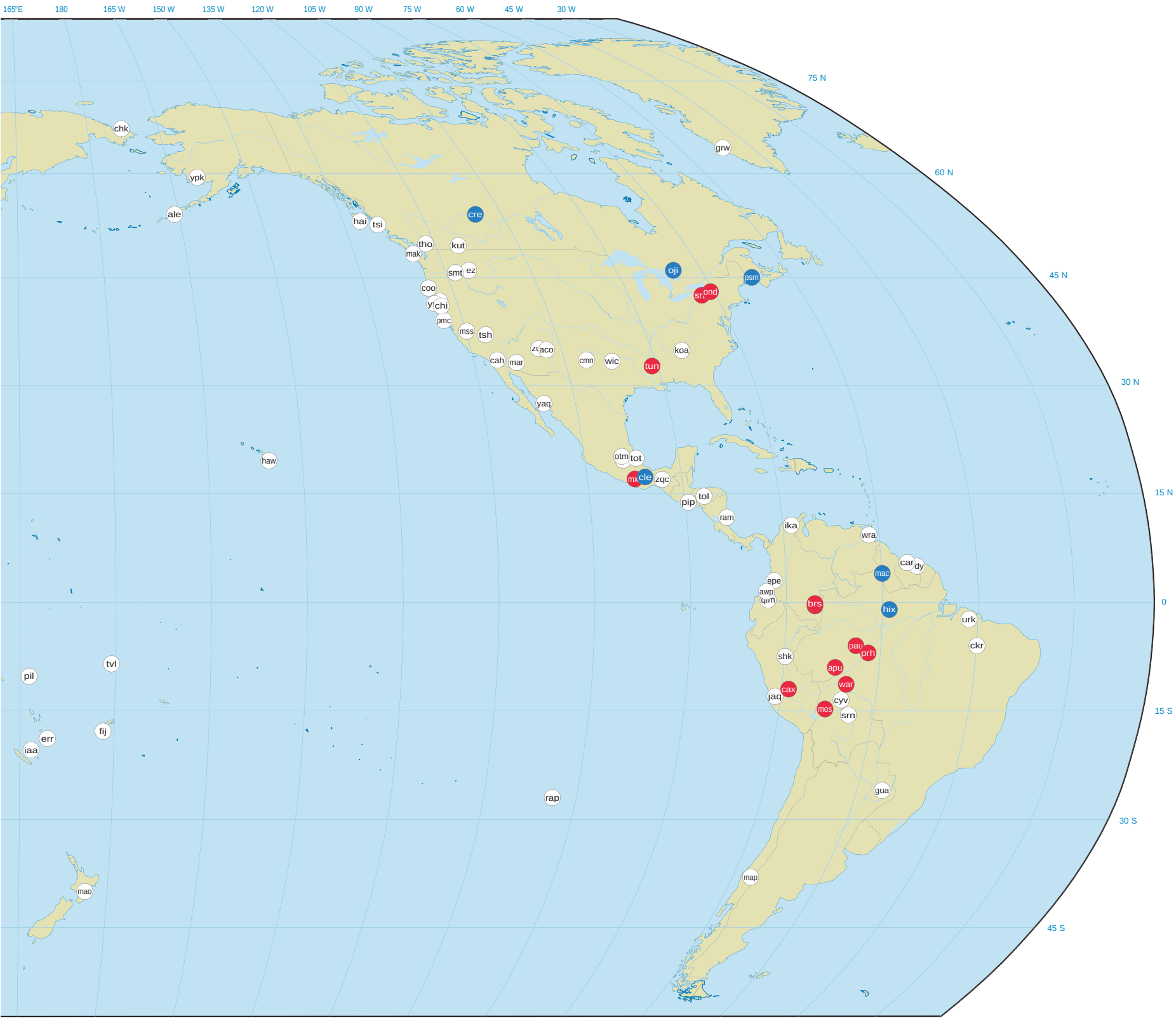


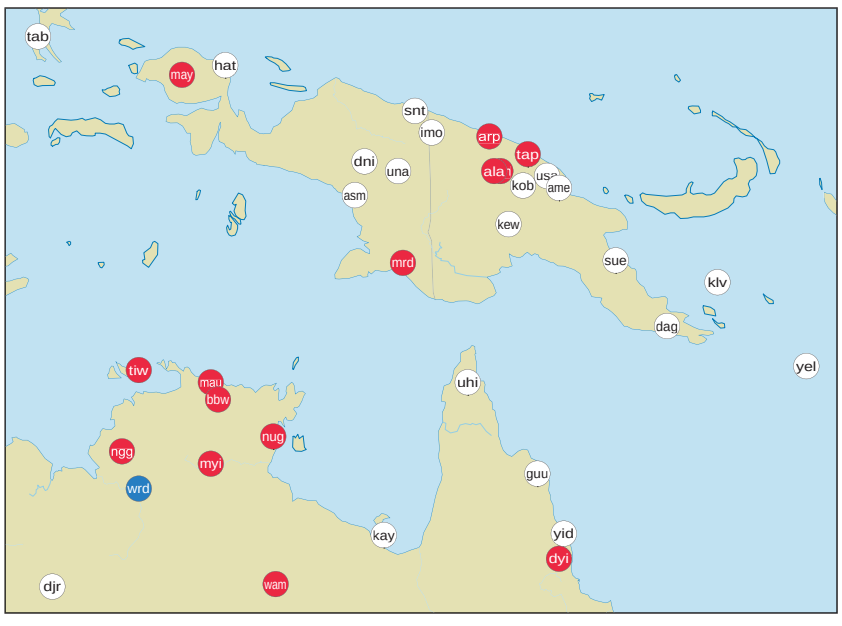


eastern Africa

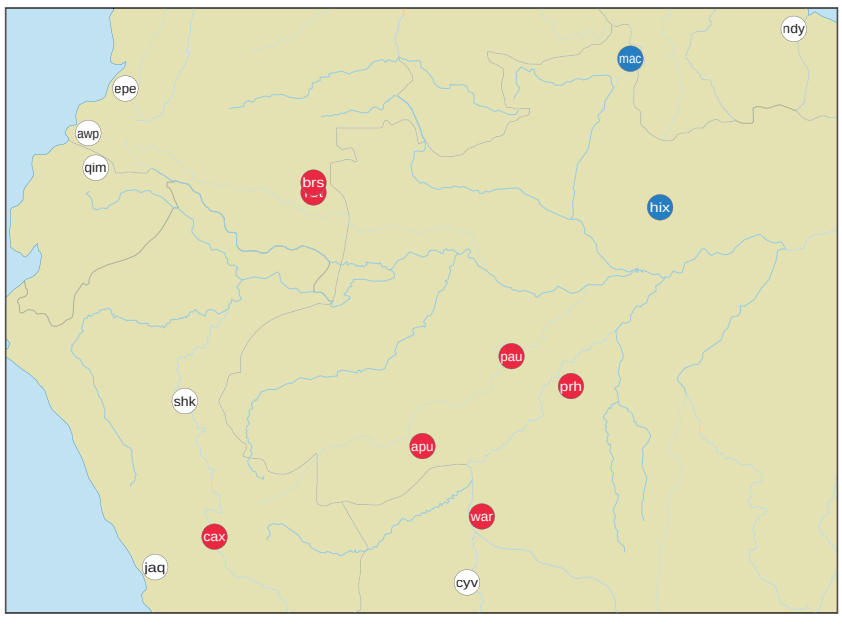


Caucasus





New Guinea and northern Australia



northwestern South America

32 Systems of Gender Assignment

GREVILLE G. CORBETT

In Chapters 30 and 31 we have seen how we as linguists analyse gender systems, and establish how many genders there are. The remaining question is how the *speaker* assigns nouns to those genders. In other words, if a speaker of Russian uses the word *kniga* 'book' or *djadja* 'uncle', how does he or she "know" the gender? Clearly speakers must know the gender in order to be able to make the appropriate agreements. A model of the way in which speakers allot nouns to genders is called a *gender assignment system*.

1 Defining the values

Assignment may depend on two sorts of information about the noun: its meaning and its form. We start with what we shall call **strict semantic systems.** In some languages the meaning of a noun is sufficient to determine its gender, for all or almost all nouns. This type is found in Dravidian languages like Kannada (Karnataka, southern India; Sridhar 1990: 198). In Kannada, nouns denoting male humans are masculine, and those denoting female humans are feminine. There are also deities, demons, and heavenly bodies in these genders. All remaining nouns, including those denoting infants and animals, are neuter. Thus *appa* 'father' and *candra* 'moon' are masculine, *amma* 'mother' is feminine, and *na:yi* 'dog' is neuter.

Many languages have semantic assignment rules which do not cover the noun inventory as completely as do the rules of Kannada. We shall call these **predominantly semantic assignment systems**. An example is found in Bininj Gun-Wok, which was introduced at the beginning of Chapter 31. The semantic categories found in each gender are given in Table 1.

This table is from Evans et al. (2002), and full details of gender assignment in Bininj Gun-Wok can be found there; the items in square brackets represent categories which have moved into the vegetable gender from the neuter in the speech of younger speakers. The important points for our typology are that the semantic assignment rules are considerably more complex than those of Kannada, and yet the coverage is not as good. For example, lower animates are split between the masculine and feminine genders, and it is hard to be more specific; nouns denoting reptiles, birds, and fish are found in both categories. There may well be principles of categorization here of which we are still unaware, but it seems likely that for at least some nouns there is no longer a principle for assignment which is still "live" for current speakers.

The genders of Bininj Gun-Wok have a semantic core, like those of Kannada. However, the rules of Bininj Gun-Wok are more complex, and still leave more nouns unaccounted for. This typological distinction applies equally well to languages where the dominant semantic principle involves animacy rather than sex. We noted in Chapter 31 how Eastern Ojibwa (Algonquian; Ontario, Canada) assigns nouns to gender according to animacy, but how some nouns do not fall readily under the rule. Thus in languages with semantic assignment systems, the meaning of the noun determines gender. In the strict assignment systems, the rules are obvious and cover (virtually) the entire noun inventory. In the predominantly semantic systems, there is a minority of exceptions; these exceptions have been claimed to be largely only apparent in some languages, once the cultural setting of the language is taken into account. When we ask which are the semantic criteria on which semantic systems can be based, we see recurring patterns and occasional surprises (Corbett 1991: 30–2). For the present map we treat strict semantic and predominantly semantic systems together.

In many languages, however, assignment by semantic rules would leave many nouns without an assignment to a gender. In languages like Kannada, the nouns not assigned by the semantic rules (the "remainder" or "semantic residue") all belong to a single gender. In the languages we consider next, these residue nouns are distributed over more than one gender. Here we find additional rules for assigning nouns to genders according to their form. There is a significant asymmetry: languages may base their assignment system on semantic rules, or on semantic *and* formal rules, but not just on formal rules. Formal assignment rules may in turn access two types of information: phonological and morphological. There may be combinations of such rules. We shall take a clear instance of each, considering languages from the sample.

A good example of assignment depending on phonological information is provided by Qafar (East Cushitic; northeastern Ethiopia and Djibouti; Parker and Hayward 1985). In Qafar the semantic assignment rules are fairly standard, namely, for sex-differentiable nouns, those denoting males are masculine and those denoting females are feminine. It is the nouns which fall outside these semantic rules, the residue, which are of interest. For them there are the following phonological assignment rules: nouns whose citation form ends in an accented vowel are feminine (for example, *karmà* 'autumn'), while all others are masculine (for example, *gilàl* 'winter', which does not end in a vowel, and *tàmu* 'taste', which does end in a vowel, but not an accented one). These rules operate with few exceptions. Moreover, nouns denoting males and females typically accord with them too (for example, *bàqla* 'husband' and *barrà* 'woman, wife'). It might seem that we could dispense with semantic rules for Qafar. However, while the phonological rules give the right result in almost all cases, there are some nouns which show the role of the semantic rules. We find *abbà* 'father', which is masculine, even though it ends in an accented vowel. Conversely, *gabbixeèra* 'slender-waisted female' is feminine, though the accent is nonfinal. In such cases of conflict, the semantic rules take precedence (as is the normal situation in gender assignment systems). Qafar has remarkably simple phonological assignment rules, which assign semantically heterogeneous nouns to the appropriate gender by reference to their form.

The second type of formal assignment rule accesses morphological information. Here Russian is a good example. Once again for

Table 1 The semantics of gender in Bininj Gun-Wok

Masculine	Feminine	Vegetable	Neuter
• Male higher animates • Overall default for animates • Some lower animates • Rain • Compass points • Some items used in painting • Trade items, esp. Macassan and European • Some types of honey	• Female higher animates • Some lower animates • Sun	• Plants and their products, including life-form terms • Sexual and excretory body parts • Song, ceremony, and custom • Fire (both bush and domestic) • Food, vegetable and otherwise • Some types of honey • Boats, planes, and cars • [Drink, water, well] • [Camp nexus] • [Landscape features with water associations]	• Most parts of animals and plants • Some parts of the landscape • Weather and sea • Time measures • Languages and speech • Country; place-based social categories

sex-differentiables, nouns denoting males are masculine and those denoting females are feminine. But unlike the situation in languages like Kannada, the residue is shared between the three genders, with the neuter gender not even receiving the majority. We might think that further semantic rules would be sufficient, but this turns out to be at best highly unlikely; see the data in Table 2, where the nouns in each row are semantically similar yet belong to three different genders.

Table 2 Russian nouns belonging to the semantic residue

masculine		feminine		neuter	
žurnal	'magazine'	*gazeta*	'newspaper'	*pis'mo*	'letter'
stul	'chair'	*taburetka*	'stool'	*kreslo*	'armchair'
dom	'house'	*izba*	'hut'	*zdanie*	'building'
čaj	'tea'	*voda*	'water'	*vino*	'wine'
ogon'	'fire'	*peč'*	'stove'	*plamja*	'flame'
dub	'oak'	*bereza*	'birch'	*derevo*	'tree'
avtomobil'	'car'	*mašina*	'car'	*taksi*	'taxi'
flag	'flag'	*èmblema*	'emblem'	*znamja*	'banner'
glaz	'eye'	*ščeka*	'cheek'	*uxo*	'ear'
lokot'	'elbow'	*lodyška*	'ankle'	*koleno*	'knee'
nerv	'nerve'	*kost'*	'bone'	*serdce*	'heart'
večer	'evening'	*noč'*	'night'	*utro*	'morning'
čas	'hour'	*minuta*	'minute'	*vremja*	'time'

Thus the nouns of the semantic residue are scattered across the three genders in Russian. This situation is presented in Table 3.

Table 3 Gender assignment in Russian

masculine	feminine	neuter
Sex differentiables denoting males	Sex differentiables denoting females	Part of semantic residue
PLUS part of semantic residue	PLUS part of semantic residue	

In order to see how the remaining nouns are assigned, rather than looking at their meaning we should look instead at their morphology. There are four main inflectional classes in Russian, each with several thousands of nouns (for justification of this view see Corbett 1982: 202–11). There are six cases and two numbers (though no paradigm has twelve distinct forms because of various syncretisms). We give just the singular forms in Table 4.

Table 4 Inflectional classes in Russian

	I	II	III	IV
Nominative	*žurnal*	*gazeta*	*kost'*	*pis'mo*
Accusative	*žurnal*	*gazetu*	*kost'*	*pis'mo*
Genitive	*žurnala*	*gazety*	*kosti*	*pis'ma*
Dative	*žurnalu*	*gazete*	*kosti*	*pis'mu*
Instrumental	*žurnalom*	*gazetoj*	*kost'ju*	*pis'mom*
Locative	*žurnale*	*gazete*	*kosti*	*pis'me*
gloss	'magazine'	'newspaper'	'bone'	'letter'

Given information about the inflectional class of nouns, the assignment rules are straightforward. Nouns in class I are masculine, those in classes II and III are feminine, and those in IV are neuter. (Further rules are required in Russian for indeclinable nouns, like *taksi* 'taxi', which is neuter; however, indeclinability is itself a morphological property.) In view of the coverage of these rules, we might be tempted to think that we could dispense with the semantic assignment, since *mal'čik* 'boy' is in class I, while *sestra* 'sister' is in class II, and *mat'* 'mother' is in class III. In other words, many of the sex-differentiable nouns would be assigned to the appropriate gender by the morphological assignment rules. But there are also instances where this is not so, for instance, *djadja* 'uncle', which denotes a male but is in class II, whose nouns are typically feminine. *Djadja* 'uncle' is masculine. Nouns like this confirm, once again, that we do not find languages where formal assignment rules are sufficient.

Of course, there are languages where the rules are more complex than these, but for languages where careful research has been undertaken, gender is always predictable from a set of assignment rules, for at least 85 per cent of the noun inventory and usually for a substantially larger proportion than that. For more details on these assignment systems see Corbett (1991: 7–69). For the purposes of the map we shall treat phonological and morphological assignment together. Thus the division will be between languages which have semantic assignment rules on the one hand, and those which have both semantic and formal assignment rules on the other. The values are as follows:

○	1. No gender system	144
●	2. Semantic assignment	53
●	3. Semantic and formal assignment	59
	total	256

2 Geographical distribution

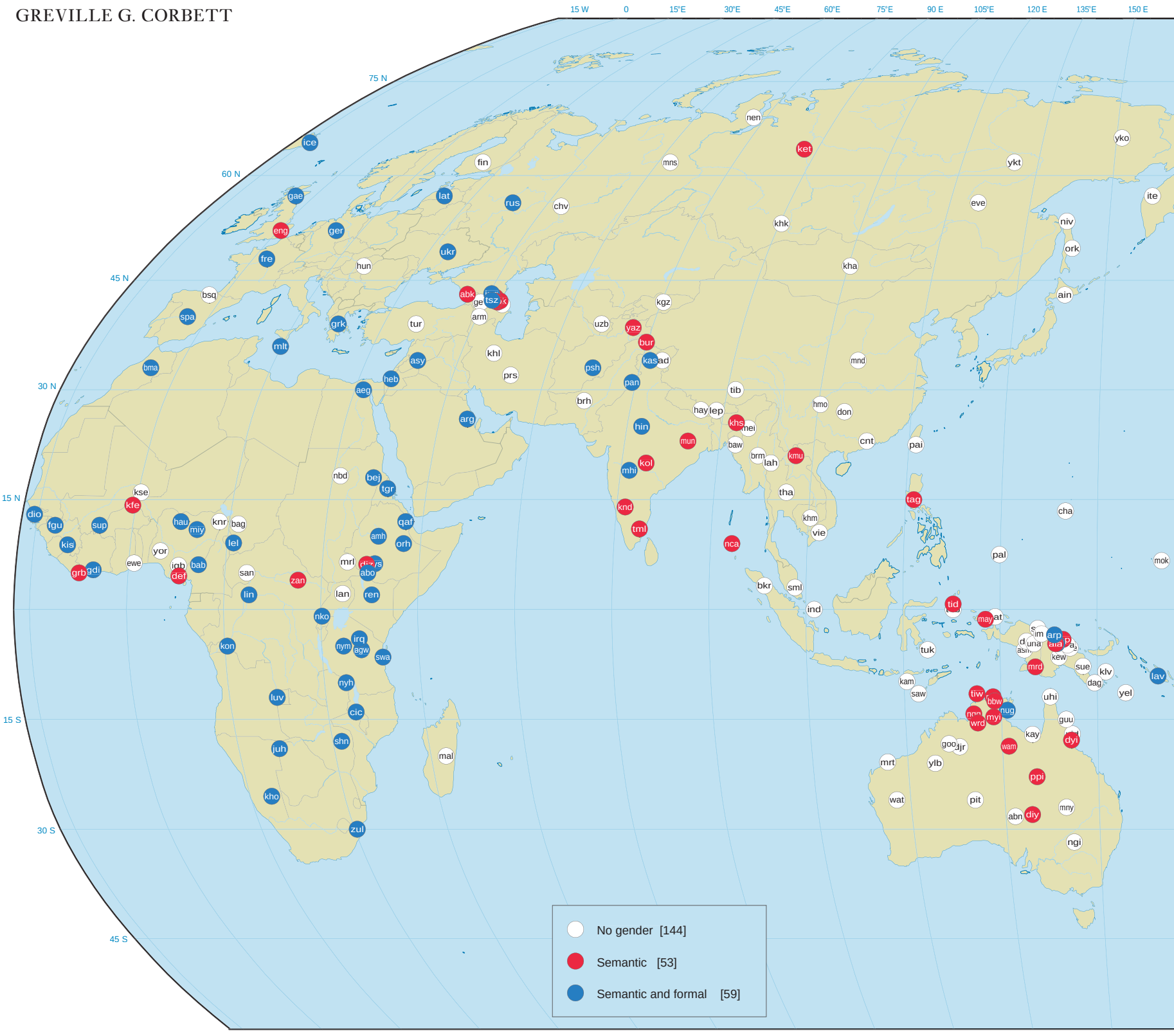
Of the languages in our sample with gender systems, there is a roughly even split between the two types of assignment system: 53 have semantic assignment while 59 have semantic and formal assignment. The distribution is interesting. Semantic and formal assignment is found mainly in Eurasia and Africa, in the Indo-European, Afro-Asiatic and Niger-Congo families. The convincing accounts of the rise of gender systems provide paths leading to systems with semantic assignment. And there are ways in which such systems may further develop into systems with semantic and formal assignment. Note that there is no necessity for this development to occur: the Dravidian systems have remained as semantic systems for a substantial period. However, one might reasonably expect that if a system is of the semantic and formal assignment type, this is likely to indicate an "old" gender system, since there must have been sufficient time for the system to develop from an earlier semantic assignment system. By contrast, if the system is of the predominantly semantic type we can make no prediction, as such systems can arise at any time. It would follow that "old" gender systems are found in Eurasia and Africa.

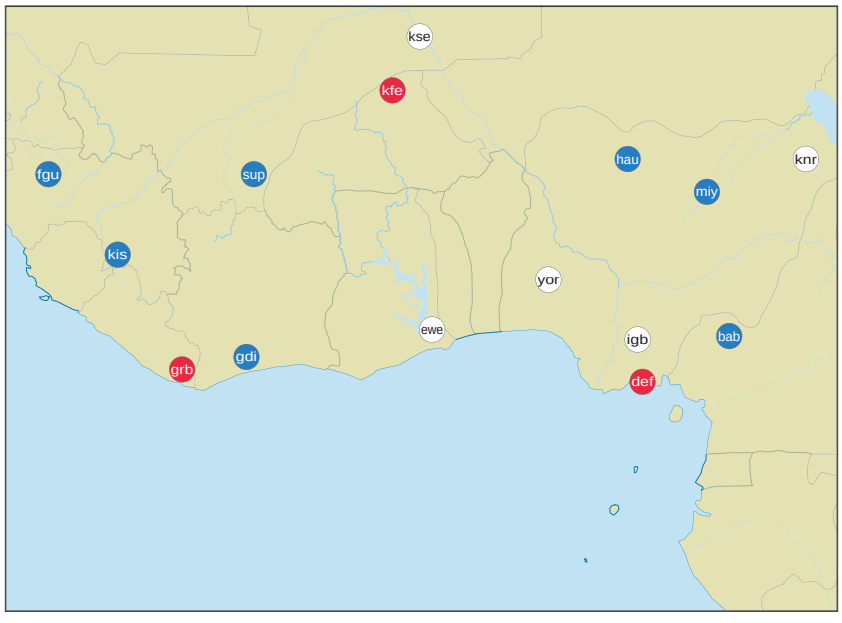
3 Theoretical issues

There are several theoretical issues which arise. Given the typology of assignment systems, it is natural to ask what other features might be correlated with the assignment system. There have been several instances of modelling assignment systems, particularly those of the more difficult types, where it is helpful to be able to demonstrate that the proposed system does indeed account for the vast majority of the nouns in the lexicon (see Fraser and Corbett (1995) on Russian, Fraser and Corbett (1997) on Arapesh, both discussed in Corbett and Fraser (2000), and Evans et al. (2002) on Bininj Gun-Wok). Since it has been shown that gender is always largely predictable, this raises an interesting issue for lexicologists: what is the status of a lexical feature which is predictable? Psycholinguists too are beginning to tackle the issue of the place of gender in lexical entries. There has been some interesting work on how children acquire gender systems, for example, Mills (1986) and Müller (2000). Such studies may also help us to see how such systems change over time, as shown by the work of Polinsky and Jackson (1999), on Tsez; see also Comrie and Polinsky (1998); for development of the work on modelling change in assignment systems see Polinsky and van Everbroeck (2003). There are interestingly different gender systems to investigate, and so it is important that we are careful about definitions, in order to ensure that our comparisons are valid.

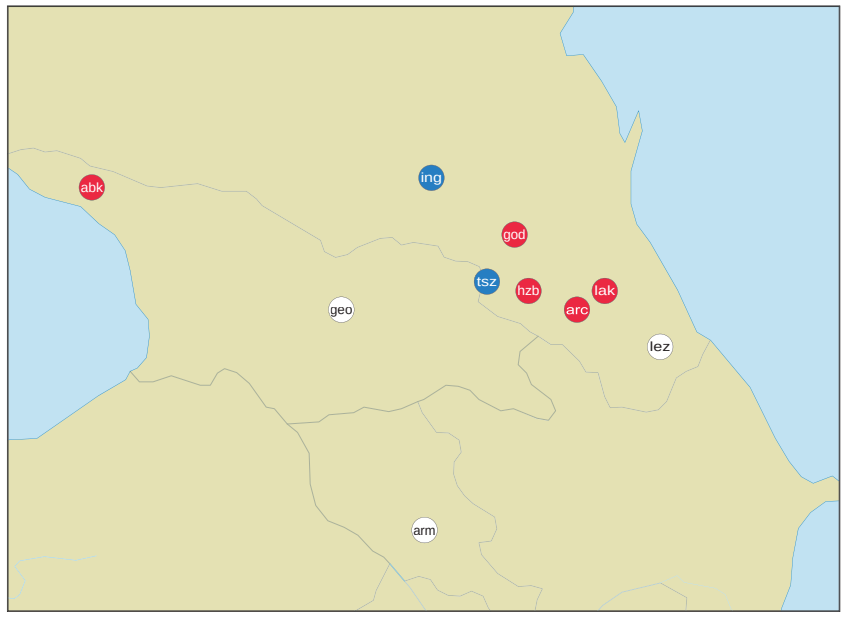
32 Systems of Gender Assignment

GREVILLE G. CORBETT

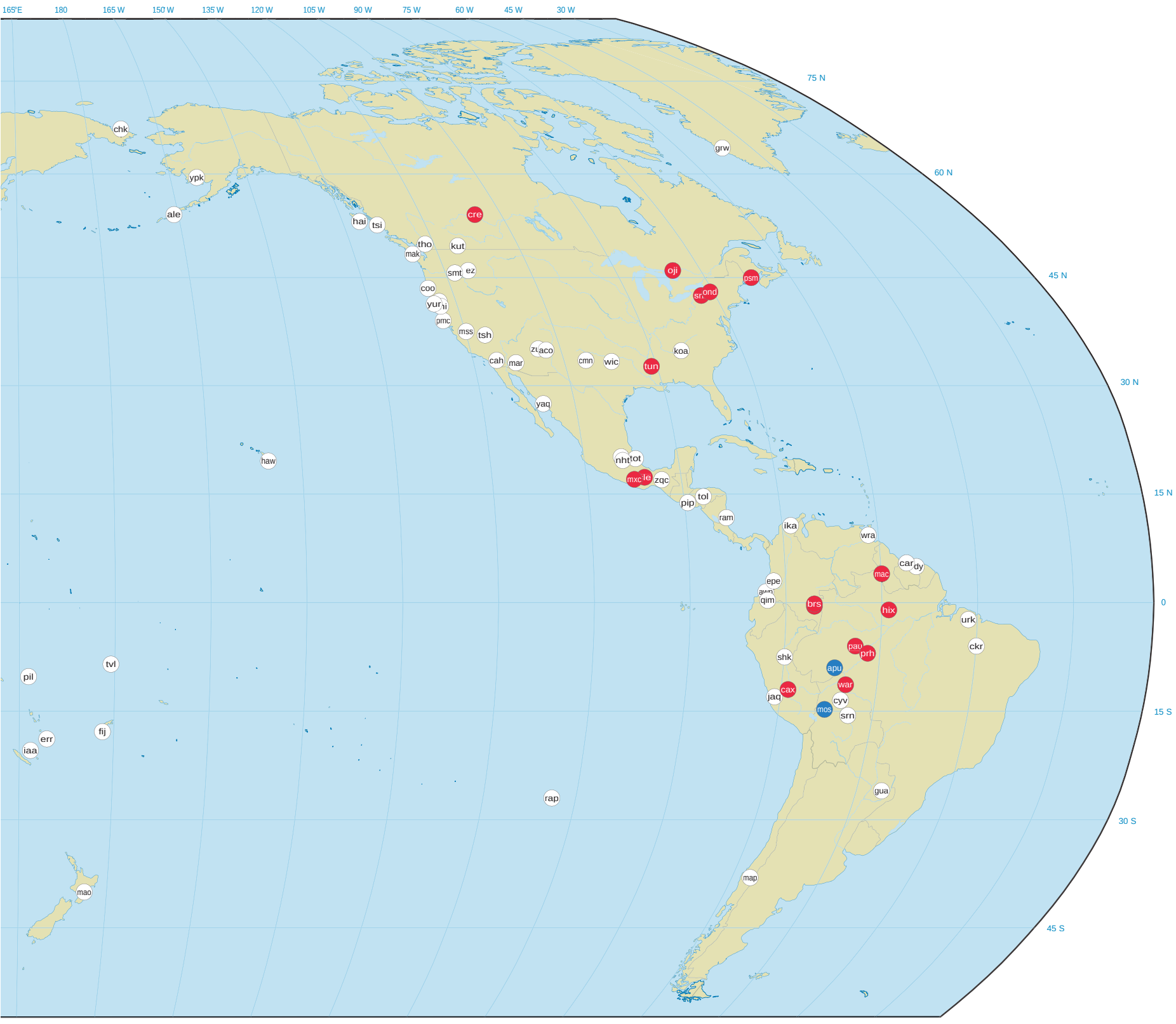


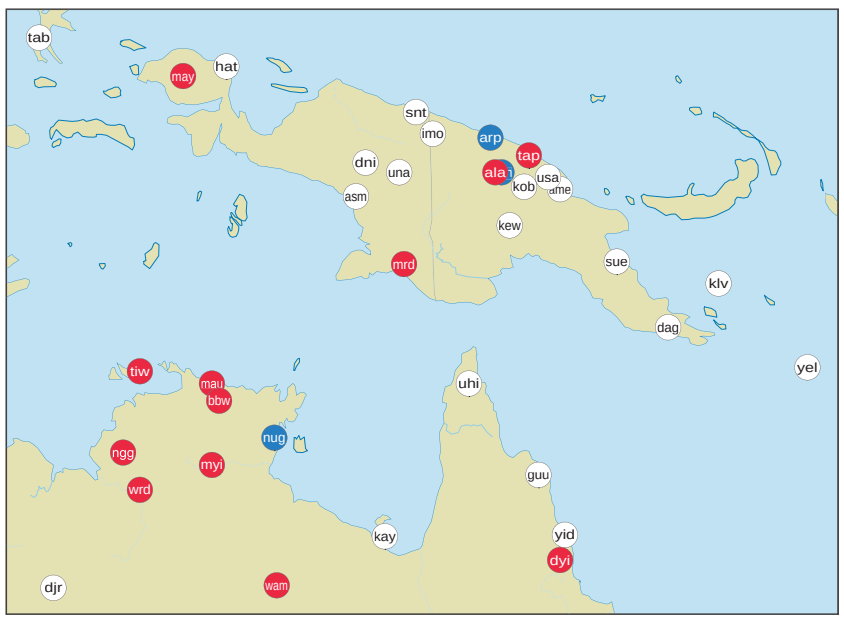


West Africa

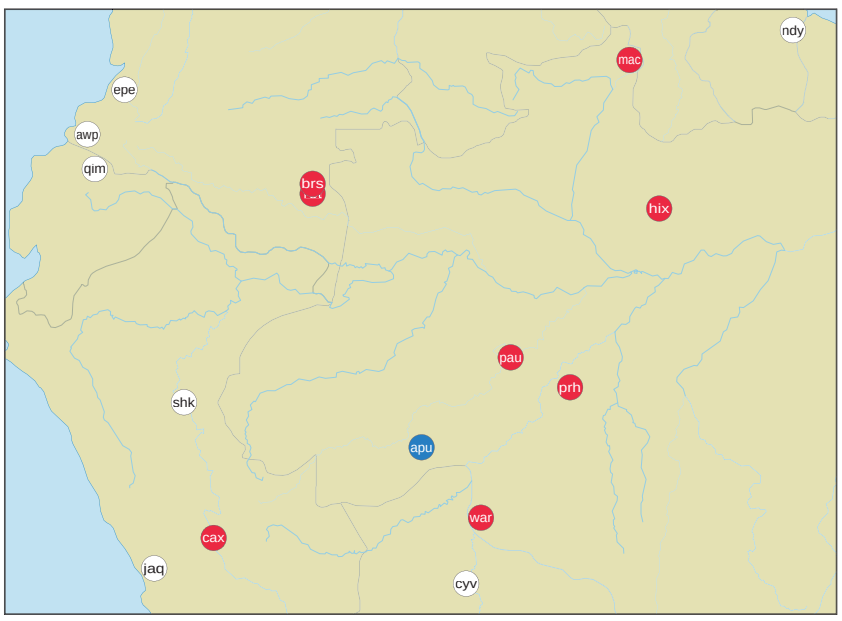


Caucasus





New Guinea and northern Australia



northwestern South America

33 Coding of Nominal Plurality

MATTHEW S. DRYER

1 Defining the values

This map shows the method by which a language indicates plurality with nouns. There are basically two ways in which languages indicate plurality. The first (and most common) involves changing the morphological form of the noun, as in English *dog, dogs*. The second involves indicating plurality by means of a morpheme that occurs somewhere else in the noun phrase, illustrated by the plural word in the example in (1) from Hawaiian, where the word *mau* has the same function as the plural suffix in English, but is a separate word modifying the noun.

(1) Hawaiian (Elbert and Pukui 1979: 159)
ʻelua *aʻu* *mau* *iʻa*
two my PL fish
'my two fish'

1.	Plural prefix	118
2.	Plural suffix	495
3.	Plural stem change	5
4.	Plural tone	2
5.	Plural by complete reduplication of stem	8
6.	Morphological plural with no method primary	34
7.	Plural word	150
8.	Plural clitic	59
9.	No plural	86
	total	957

Six of the values shown on the map involve ways of indicating plurality morphologically on the noun. The first of these is **plural prefixes**, illustrated by the example in (2) from Anindilyakwa (isolate; Northern Territory, Australia).

(2) Anindilyakwa (Leeding 1989: 294)
wirr-*iyikwayiwa*
PL-child
'children'

The second of these is **plural suffixes**, illustrated by English and by the example in (3) from Nagatman (isolate; Papua New Guinea).

(3) Nagatman (Campbell and Campbell 1987: 2)
*a`ma-****re***
dog-PL
'dogs'

The next three ways of forming plurals are considerably less common than the first two. Five languages are shown on the map as using **changes within the noun stem** as the primary means of forming plurals. These are two Nilotic languages in Sudan, Dinka (Nebel 1948: 3) and Nuer (Crazzolara 1933: 28; Frank 1999), two Yuman languages of the southwestern United States, Maricopa (Gordon 1986: 29) and Jamul Tiipay (Miller 2001: 115), and Laal, an isolate spoken in Chad (Boyeldieu 1982: 70–82). The primary means for forming plurals in these languages involves changes to the vowels within the noun stem, a means used for a number of nouns in English with irregular plurals (e.g. *man, men; goose, geese*). The examples in (4) illustrate this for Maricopa.

(4) Maricopa (Gordon 1986: 29)

humar	'child'	*humaar*	'children'
nchen	'older sibling'	*nchiin*	'older siblings'
hat	'dog'	*haat*	'dogs'
mhay	'boy'	*mhaa*	'boys'

Note that Dinka and Nuer use stem changes as the primary means for forming different case forms as well (see Map 51).

The fourth way to form plurals is by **tone**. This is the primary means for only two languages shown on the map, Gworok (Niger-Congo; Nigeria; Adwiraah 1989: 28–32) and Ngiti (Central Sudanic, Nilo-Saharan; Democratic Republic of Congo), the latter illustrated in (5).

(5) Ngiti (Kutsch Lojenga 1994: 135)

kamà	'chief'	*kámá*	'chiefs'
màlàyikà	'angel'	*màlàyíká*	'angels'
màlimò	'teacher'	*màlímó*	'teachers'
adɔ̀du	'my brother'	*adɔ́du*	'my brothers'

The fifth way to form plurals is by **complete reduplication** of the noun stem. Complete reduplication is the primary method for expressing plurality in eight languages on the map. In Indonesian, for example, the plural of *rumah* 'house' is *rumah-rumah* 'houses', and the plural of *perubahan* 'change' is *perubahan-perubahan* 'changes' (Sneddon 1996: 16–17). Note that if a language expresses plurality by reduplicating only part of the stem, then the language is classified here as prefixing or suffixing depending on whether it is the initial segment or the final segment of the stem that is reduplicated. For example, the normal means for pluralizing nouns in Pipil (Uto-Aztecan; El Salvador) is by reduplication of the first syllable of the noun stem as a prefix followed by /h/ (e.g. *rayis* 'root', *rahrayis* 'roots'; *tukat* 'spider', *tuhtukat* 'spiders'; Campbell 1985: 52); this is treated as prefixation for the purposes of this map.

The final type of languages that form their plurals by morphological means (the sixth type in the box) consists of languages that employ **two or more of the five methods just described** with little basis for saying that one of these five methods is primary, where a method is considered primary if it is used by at least twice as many nouns as any other method. For example, there are languages like Misantla Totonac (Totonacan; Mexico), in which some nouns form their plural with prefixes, as in (6a), while others form their plural with suffixes, as in (6b).

(6) Misantla Totonac (MacKay 1999: 355, 361)

a. ***lii-****šaaluh*
PL-pot
'pots'

b. *míy-kam-****án***
2.POSS-offspring-PL
'your offspring'

There are also a number of languages in which some nouns form their plurals by prefixation or suffixation while other nouns form their plurals by means of either tone or stem changes, without either of these methods being primary. Many varieties of Arabic, for example, have a large number of nouns whose plurals are formed by stem changes and a large number whose plurals are formed by suffixation. Mamvu, which is like Ngiti in being a Central Sudanic language of the Democratic Republic of Congo, combines tone with suffixation (Vorbichler 1971: 222). Note that if a language employs more than one method, such as both plural prefixes and plural suffixes, but one method is used with at least twice as many nouns as any other method, then the language is shown on the map according to the more common type. For example, in Mono (Uto-Aztecan; California), only human nouns inflect for plural; there is one reduplicative prefix that is used on six nouns while all other human nouns take a suffix, so Mono is shown as employing suffixes (Norris 1986: 76–7).

The map does not distinguish affixes that only code plurality from affixes that code plurality in combination with other inflectional features of the noun. In Russian, for example, the plural affixes also code case and grammatical gender. In Swahili and other Bantu languages, the plural prefixes also code noun class.

In addition to the six morphological types of plurality, the map shows plural words and clitics. **Plural words** are words whose meaning is like that of plural affixes, but they are separate words, as illustrated by the Hawaiian example above in (1) and the example from Chalcatongo Mixtec (Oto-Manguean; Mexico) in (7).

(7) Chalcatongo Mixtec (Macaulay 1996: 113)
Ni-xa̋a̋=rí *kʷaʔà* *žúʔa* *káni* ***xináʔa***
COMP-buy=1 many rope long PL
'I bought many long ropes.'

Plural clitics are probably best thought of as a type of plural word, whose position is defined syntactically in that they have a specified position within the noun phrase, but which attach phonologically to whatever word happens to occur adjacent to that position within a particular noun phrase. For example, in Cayuvava (isolate; Bolivia), the plural clitic *me=* attaches to whatever is the first word in the noun phrase, as in (8).

(8) Cayuvava (Key 1967: 50)
me=rišɔ *raƀ* *iri*
PL=new paddle
'new paddles'

In Sinaugoro (Oceanic, Austronesian; Papua New Guinea), the plural clitic follows other modifiers of the noun, except for numerals, which it precedes, and attaches to whatever word precedes it; in (9), it attaches to an adjective.

(9) Sinaugoro (Kolia 1975: 124)
belema *bara=**ria*** *taulatoitoi*
python big=PL six
'six big pythons'

As with plural affixes, the elements treated here as plural words and clitics occasionally code other grammatical features of the noun phrase as well. For example, in Khmer (Mon-Khmer; Cambodia) and in Maori (Polynesian, Austronesian; New Zealand) the plural words vary for definite and indefinite (Ehrman 1970: 43; Bauer 1993: 110), and thus can equally well be described as articles which vary for number. The plural words in Kisi (Atlantic, Niger-Congo; Guinea) and Sanuma (Yanomam; Venezuela and Brazil) agree with the noun in noun class (Childs 1995: 148; Borgman 1990: 144–8).

The final type shown on the map are languages apparently **lacking a morphological plural** for which there is also no evidence from available data of plural words or clitics, though in some languages such may exist. An example of a language lacking a plural is Maranungku (Daly; Northern Territory, Australia; Tryon 1970: 12). Although such languages may simply not indicate plurality at all, the plurality of nominal referents is coded on the verb if the nominal is an argument of the verb and if the language is one that codes the number of that particular argument on the verb.

There are also languages lacking a way of expressing plurality that nevertheless have means of expressing singular or dual. For example, in Aari (Omotic, Afro-Asiatic; Ethiopia), there is a singulative suffix but no plural marking (Hayward 1990a: 444). In Kayardild (Tangkic; Queensland, Australia), there is a dual suffix and a suffix that means a lot of whatever the noun denotes, but not really a plural per se (Evans 1995: 184). In Imonda (Border family; Papua New Guinea), there are only five nouns that are marked for number, but the number marking with these nouns is a suffix that indicates that a noun is either singular or dual and whose absence indicates plural in the sense of three or more (Seiler 1985: 38–9).

Languages differ considerably in how widely they indicate plurality (see Chapter 34 for further discussion). While in many languages all count nouns have plural forms, there are also many languages in which plurality is restricted to animate nouns, or to human nouns, or to a subset of human nouns. If a language only has plural forms for a very small number of nouns, it is not shown on the map as having a morphological plural. But in some languages that are shown on the map as having a morphological plural, plural morphology may be restricted to a closed class of nouns, such as kinship nouns, as in Wambon (Trans-New Guinea; Papua, Indonesia; de Vries 1989: 36).

Similar variety is found with plural words and clitics. In Kutenai (isolate; western North America), the plural clitic is only used with kinship terms and the word meaning 'friend'. Languages lacking an indication of plurality within the noun phrase are simply the limiting case along a continuum of how widely languages code plurality. It is also the case that languages differ in the extent to which indication of plurality is required with semantically plural referents, even when a morphological form is available. In many languages, such indication is not obligatory, though again this also may vary with animacy or humanness.

2 Geographical distribution

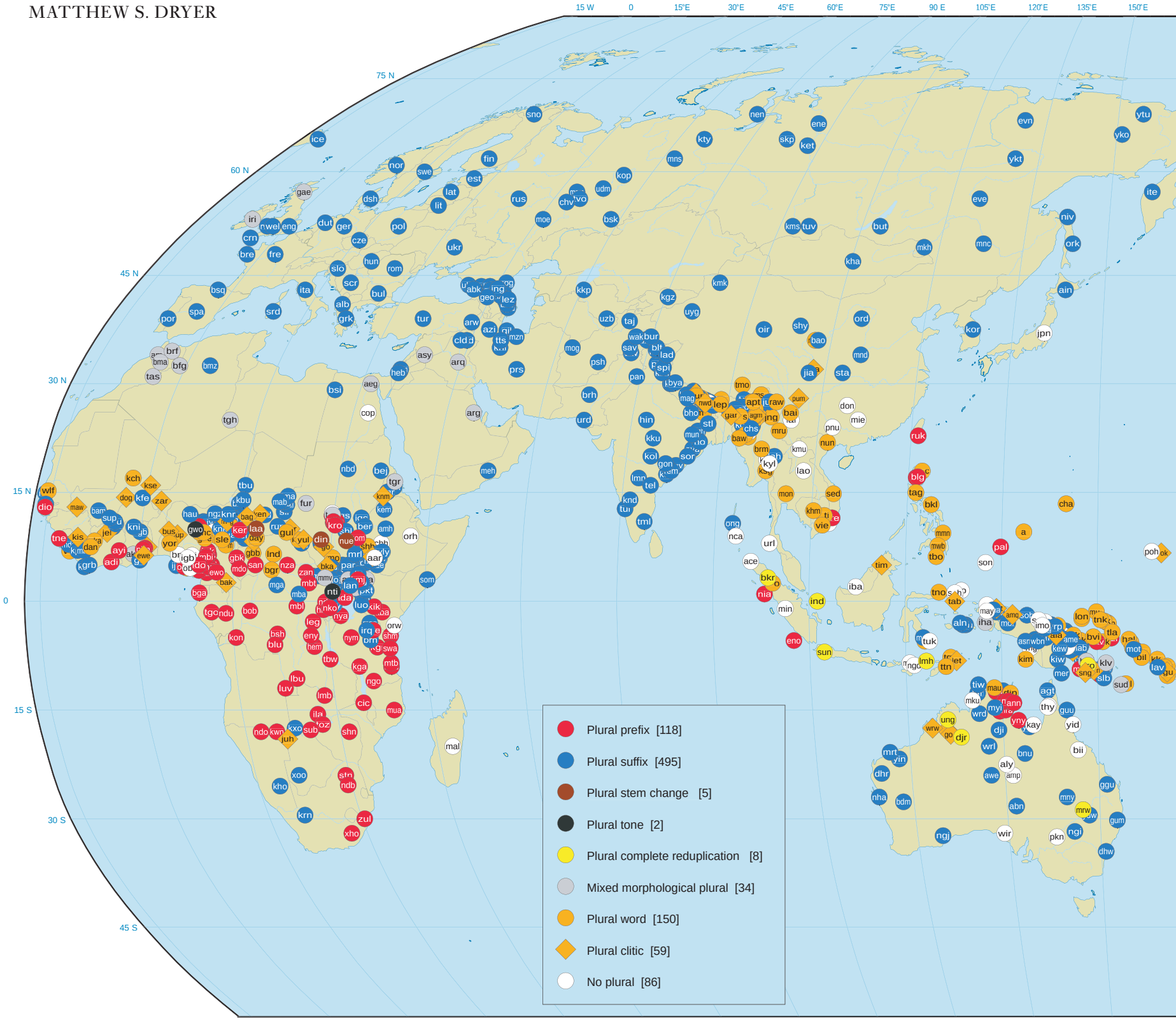
The map shows plural suffixes greatly outnumbering all other types shown on the map. They are widely distributed throughout the world; the largest area in which they are not found is South-East Asia and most of the area in which Austronesian languages are spoken. The largest area in which plural prefixes are common is in Africa, especially among Bantu languages in the southern half of the continent, but also among languages in other branches of Niger-Congo. The map shows plural prefixes scattered among Austronesian languages in the general vicinity of Indonesia and the Philippines, an area in which plural suffixes are generally lacking, but various other types are also common in this area, including plural by complete reduplication, plural words, and absence of plural altogether. Plural prefixes are completely lacking from two large areas: (i) the entire mainland of Europe and Asia, except for one language in southern Vietnam; (ii) an area in the New World stretching from Guatemala south to include all of South America. Languages with more than one primary morphological method for forming plurals are especially common in North Africa, among Berber languages and various varieties of spoken Arabic. Plural words and clitics are most common in South-East Asia and among Austronesian languages, in West Africa, and in the Amazon basin in South America.

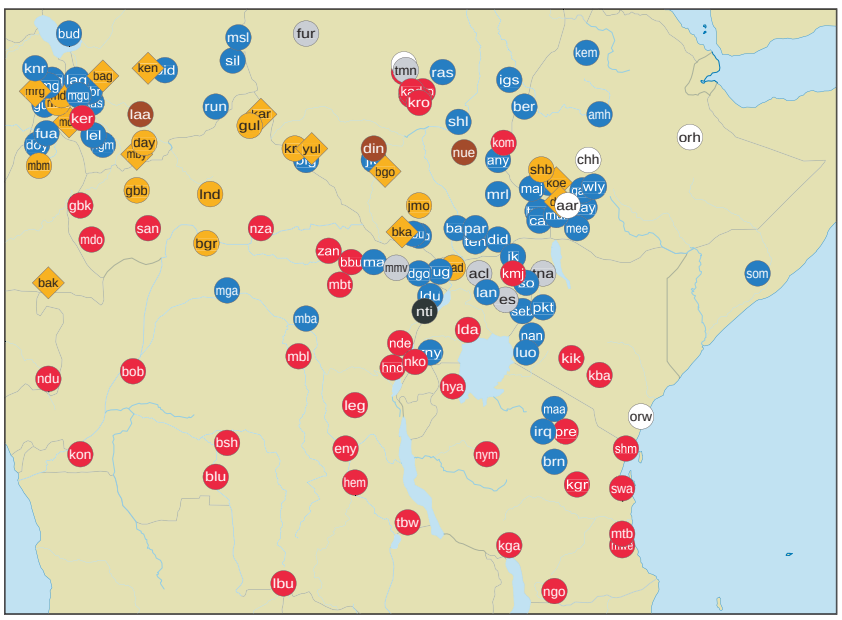
3 Theoretical issues

Corbett (2000) discusses a wide variety of theoretical issues relating to plurality and number.

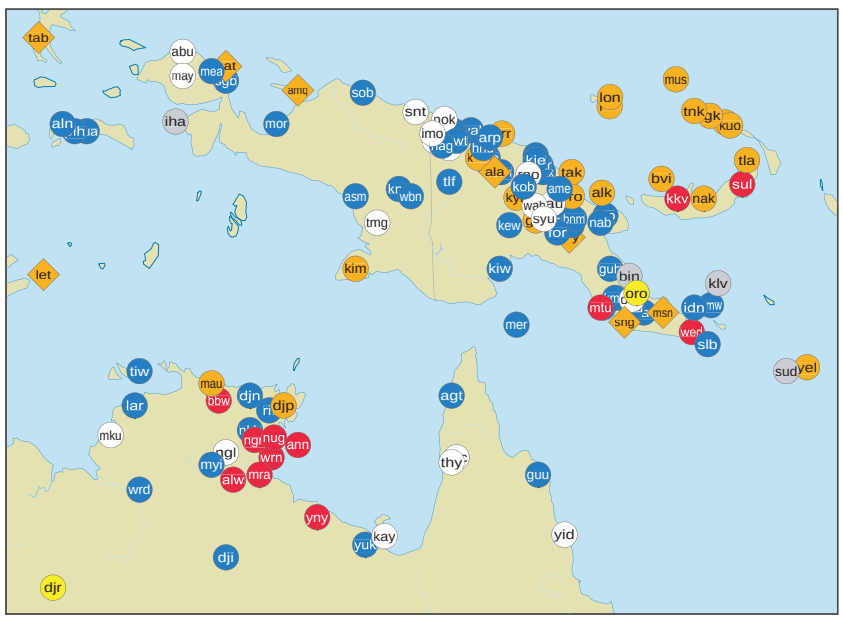
33 Coding of Nominal Plurality

MATTHEW S. DRYER

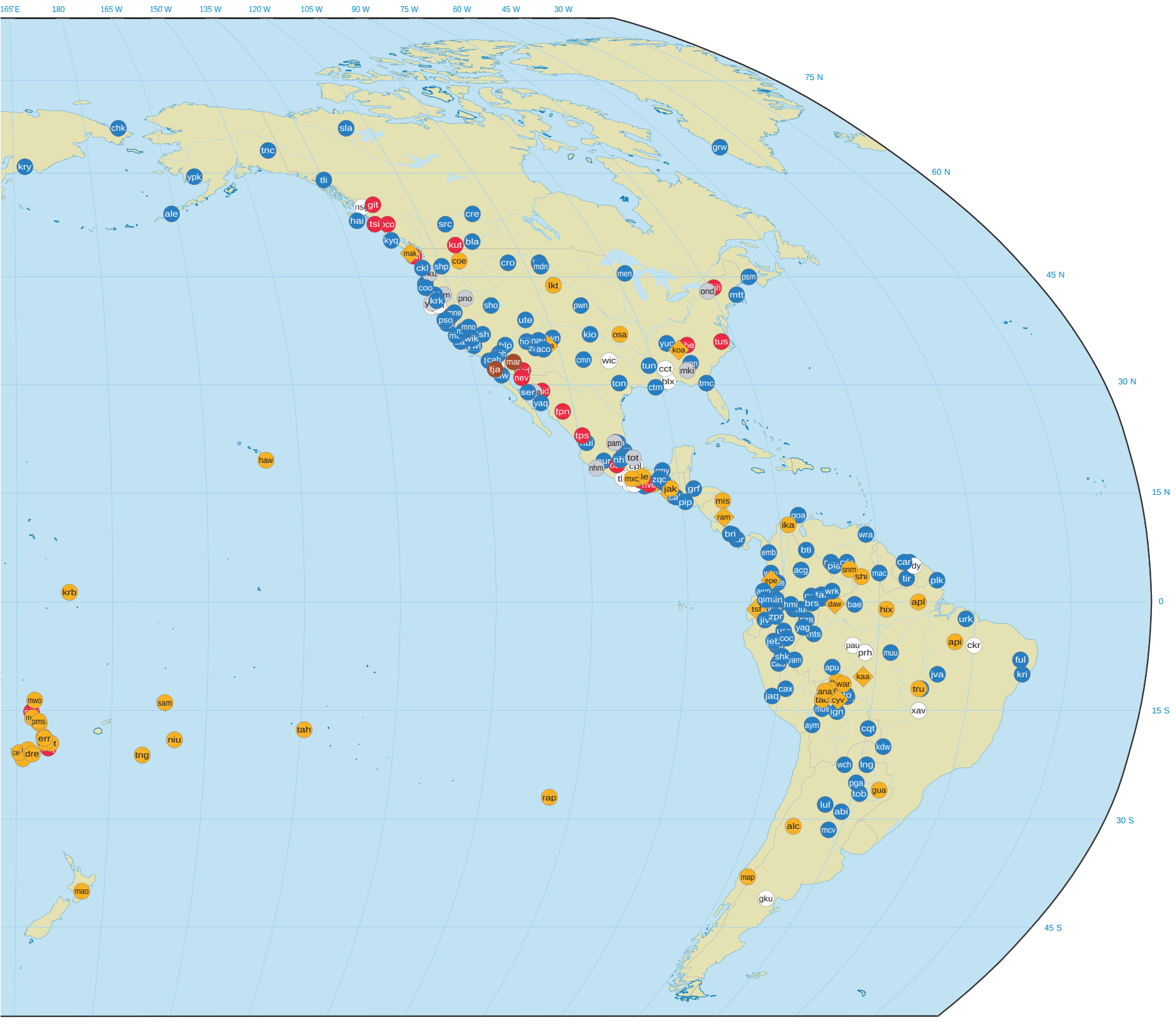




eastern Africa



New Guinea and northern Australia



western United States

Mesoamerica

34 Occurrence of Nominal Plurality

MARTIN HASPELMATH

1 Defining the values

This map shows the extent to which plural markers on full nouns are used in a language (full nouns are nouns which are not pronouns; see Chapter 35 for the plural of personal pronouns). Plural markers can have different formal manifestations (see Chapter 33), which play no role in this chapter. Here we look only at the occurrence of plural markers, which varies along at least two dimensions: animacy and obligatoriness. In the animacy dimension, the most important contrast is between **animate** (especially **human**) **nouns** and **inanimate nouns**, which are treated differently in many languages (see §2.1 for animal nouns, which are neither human nor inanimate). In the obligatoriness dimension, I distinguish here between **non-occurrence**, **optional occurrence**, and **obligatory occurrence**. When these two dimensions are combined, we get the six values shown in the feature-value box.

○	1. No nominal plural	28
○	2. Plural only in human nouns, optional	20
○	3. Plural only in human nouns, obligatory	39
○	4. Plural in all nouns, always optional	55
○	5. Plural in all nouns, optional in inanimates	15
○	6. Plural in all nouns, always obligatory	133
	total	290

An apparently exceptionless generalization is that **human nouns are more likely to have plural marking than nonhuman (especially inanimate) nouns**. Thus, three further logically possible types, in which inanimate plurals would occur more widely than human plurals, are not attested:

7. Plural only in inanimate nouns, obligatory
8. Plural only in inanimate nouns, optional
9. Plural in all nouns, optional in human nouns

It should be noted that we are talking here about the occurrence of **plural marking**, not the occurrence of **plural meaning**. Non-occurrence of plural marking does not mean that only a singular meaning can be expressed, but that a non-number-marked form of the noun is used both for a single referent and for a group of referents. For example, in Shigatse Tibetan, a language with no plural marking on full nouns, the noun *ri̱* can be translated into English as 'mountain' or 'mountains' (Haller 2000: 59, 114). While English speakers must express the number of entities in such a case, Shigatse Tibetan speakers are simply vague with respect to the number of entities (if they wish to be explicit, they can use numerals and quantity words like 'many'). Similarly, in Baka (Ubangi, Niger-Congo; Cameroon), where all nominal plurals are optional, the noun form *gba* can be translated as 'village' or 'villages'. Speakers can be explicit and use the plural form *gba-o* 'villages', but they do not have to (Kilian-Hatz 1995: 21, 95).

The first type shown on the map comprises those languages that have **no nominal plural** at all. This is not uncommon in languages with little or no inflection, but there are also quite a few languages with fairly rich inflectional systems that simply happen to lack nominal number marking. For instance, Gurr-goni (Burarran; Northern Territory, Australia) has nominal case-marking inflection and rich verbal morphology, but no noun plurals (Green 1995: 69).

Languages of the second and third type have a nominal plural **only in human nouns** and lack a plural of inanimate nouns. In the second type, this human-only plural is **optional**. An example of such a language is Hatam (West Papuan; Indonesia), where human nouns like *munggwom* 'child(ren)' may but need not take the plural clitic *=nya* when reference to several people is intended (Reesink 1999: 50–1, 65). In the third type, the human-only plural is **obligatory**. For instance, in Jamul Tiipay (Yuman; southern California) kinship terms and other human nouns have plural forms (e.g. *xechany* 'girl', *xaachaany* 'girls') which must be used, while inanimate nouns lack plurals (Miller 2001: 114–15).

Languages of the remaining three types have plural marking **in all nouns**, i.e. human nouns and inanimate nouns. In the fourth type, this marking is **always optional** (i.e. both in human and in inanimate nouns). An example of this is Tetun (Central Malayo-Polynesian; Timor), where the plural word *sia* (also used as a third-person plural pronoun 'they') can combine both with human and with nonhuman nouns (e.g. *asu* 'dog(s)', *asu sia* 'dogs'; *kaheke* 'fan(s)', *kaheke sia* 'fans'; van Klinken 1999: 124–5, 135). In the fifth type, plural marking is **obligatory in human nouns and optional in inanimate nouns**. This is exemplified by Arawak (Arawakan; Suriname), where human nouns require the plural suffix *-non* (*hiaro* 'woman', *hiaro-non* 'women') while inanimate nouns allow the plural suffix *-be* (*siba* 'stone(s)', *siba-be* 'stones'; Pet 1987: 34–7). Finally, languages of the sixth type have **obligatory plural marking in all nouns**, human and inanimate. This is exemplified by English and other European languages; another example is West Greenlandic (e.g. *nuna* 'land', *nuna-t* 'lands'; Fortescue 1984: 247).

2 Refining the definitions

The definitions of the feature values that I have just given involve two pairs of concepts that require further comment.

*2.1 **Human versus inanimate.*** For the purposes of the map, I looked only at nouns denoting humans and nouns denoting discrete inanimates. Nonhuman animates, i.e. nouns denoting (nonhuman) animals, pattern differently in different languages. In some languages, animals are treated like humans; an example is Korku (Munda; central India): *siṭa/siṭa-ku* 'dog/dogs', like *koro/kor-ku* 'man/men', as opposed to pluralless inanimates (Nagaraja 1999: 31, 35). In other languages, animals are treated like inanimates; an example is Taba (Eastern Malayo-Polynesian; Maluku, Indonesia): *manik* 'chicken(s)', like *llu* 'leaf/leaves', as opposed to human nouns with obligatory plural marking, e.g. *wang/wang=si* 'child/children' (Bowden 2001: 190–1).

So this is an interesting parameter of variation, and the main reason animal nouns are disregarded here is that many grammars provide no information about them. In some languages, the class of animals itself is subdivided further, with "higher" animals patterning like humans and "lower" animals patterning like inanimates. For instance, in Tiwi (Tiwian; Northern Territory, Australia), three animal names ('dog', 'dingo', and 'goanna') have plurals like human nouns (Osborne 1974: 52).

In some languages, not all human nouns are treated alike. For instance, in Amele (Trans-New Guinea; Papua New Guinea), only kinship terms have plurals (e.g. *cot-i* 'my brother', *cot-i-el* 'my brothers'; Roberts 1987: 171). In other languages, the subset of human nouns with plurals does not fall under any semantic generalization and is purely lexically determined. For the purposes of the map, such cases were not distinguished from languages where all human nouns have plurals.

Finally, in many languages not all inanimates are treated alike: nouns denoting nondiscrete masses such as 'sand', 'salt', 'water', 'milk' often lack plurals even in languages that otherwise obligatorily mark the plural of inanimates (e.g. in English). It is possible that this case is somewhat different from the non-occurrence of plurals in discrete nouns, because English speakers find it difficult to imagine what the semantic difference between *milk* and *milks* would be,

especially if we leave aside the sort reading ('different sorts of milk'). However, there are quite a few languages in which speakers find it easy to form and interpret plurals of nondiscrete nouns, for instance in Evenki (Tungusic; eastern Siberia), where such plurals are said to denote a large quantity or several pieces (e.g. *ulle* 'meat', *ulle-l* 'a lot of meat, pieces of meat'; Nedjalkov 1997: 190).

Thus, it is likely that a more fine-grained study of plurals would find considerable evidence for a richer implicational scale as in (1).

(1) kinship terms > other humans > "higher" animals > "lower" animals > discrete inanimates > nondiscrete inanimates

This scale should be read as follows: if a noun subclass has plurals, then so do all subclasses to its left on the scale. And if a noun class has obligatory plural marking, then so do all noun subclasses to its left on the scale.

2.2 Obligatory versus optional. In principle, obligatory occurrence of plural marking means that whenever plural reference is intended, the plural must be used. However, languages with otherwise obligatory plural marking very often allow it to be absent when the noun co-occurs with numerals and other quantity words. For instance, in Lezgian (Daghestanian; eastern Caucasus), the plural is normally obligatory, but it is impossible when the noun is accompanied by a numeral (e.g. *nük'-er* 'birds', but *c'ud nük'* 'ten birds', not **c'ud nük'-er*, Haspelmath 1993: 232). In Itzaj (Mayan; Guatemala), the plural is normally obligatory, but it is optional with numerals (Hofling 2000: 227). In Brahui (Dravidian; western Pakistan), the plural is normally obligatory, but is not used with numerals in indefinite NPs (e.g. *irā bandagh(*-āk)* 'two persons', but *'amē bīstangā bandagh-āk* 'those twenty persons'; Andronov 1980: 34–6). When a grammar says that the plural is not used when a quantity expression is present, this is not regarded as evidence for general optionality of the plural; in fact, it can be taken as indirect evidence for obligatoriness, because if the plural were genuinely optional, no special statement about quantity expressions would be necessary. So, like animacy, the dimension of obligatoriness consists of more distinctions than are made here, but since many grammars do not say what happens when a noun is combined with a numeral, it was necessary to simplify the picture and loosen the criteria for obligatoriness. Another special circumstance under which plural marking is often lacking is nonreferential or generic use. For instance, Evenki allows singular nouns in contexts like (2).

(2) Evenki (Nedjalkov 1997: 190)
Bi *uluki-je* *va:-d'a-m.*
I squirrel-INDEF.ACC kill-PRES-1SG.SUBJ
'I hunt for squirrels.'

Again, when a grammar mentions optionality of plural marking in such contexts, this is not taken as evidence for genuine optionality here.

Plural marking is classified as optional when it is explicitly described as being optional, and if this optionality clearly goes beyond the non-occurrence of plurals in the two special contexts just mentioned. In many cases, it is probably possible to identify further conditions on the use of plurals. For instance, after noting that the plural word *àwọn* is optional in Yoruba, Rowlands (1969: 41) goes on to explain that the plural marker is used if the referent is thought of as a group of individuals, whereas no marker is used if it is thought of collectively (e.g. *àwọn ìwé mi* [PL book my] 'my (various) books', versus *ìwé mi* [book my] 'my (collection of) books' or 'my book'). For Udihe (Tungusic; eastern Siberia), Nikolaeva and Tolskaya (2001: 116) observe: "The Plural of a subject, especially when denoting a person, is more or less regularly marked, whereas the Plural suffix of a direct object appears less frequently." Occasionally, the conditions are described in a very precise way. In Kâte (Finisterre-Huon; Papua New Guinea), a human noun can (and indeed must) have plural marking only if it is possessed (e.g. *motec* 'boy(s)', *motec-fâc-ticne* [boy-PL-3SG.POSS] 'his boys'; Pilhofer 1933: 43, 56, 113). In this case, there is strictly speaking no optionality; rather, the plural suffix *-fâc* is obligatory when the noun is possessed and obligatorily absent when it is not possessed. Still, whenever the variable appearance of plural marking depends on a set of factors unrelated to quantity expressions and nonreferentiality, I classify the language as showing optional plural marking, because it would not be possible to systematically distinguish subtypes of optionality.

Unfortunately, a large number of grammatical descriptions do not say whether plural marking is obligatory or optional. I suspect that in most such cases, the absence of a statement about obligatoriness is due to the fact that plural marking is obligatory in the language just as in the European languages that the readers are likely to be familiar with. Authors of such descriptions may not even be aware that obligatoriness of plural marking is an issue in many languages. Grammar writers are much less likely to omit a statement about obligatoriness/optionality if plural marking is optional, because in this case their language is interestingly different from the European languages. For this reason, I decided to take obligatoriness of plural marking as the default situation and to classify a language as obligatorily plural-marking even in the absence of an explicit statement, unless the description is very sketchy or there are other reasons for suspecting that plural marking may not be obligatory (if this was the case, I excluded the language from the sample). It is possible that this decision introduces a bias into the classification, and that some languages classified as obligatorily plural-marking have in fact only optional plural marking.

3 Geographical distribution

Obligatory plural marking of all nouns is found throughout western and northern Eurasia and in most parts of Africa. The rest of the world presents a heterogeneous picture. Optional plural marking is particularly common in South-East and East Asia and Australia, and complete lack of plural marking is particularly found in New Guinea and Australia. In addition to the areal correlations, there also seems to be at least one correlation with the general morphological type: isolating languages (i.e. languages with little inflectional affixation; see Chapter 22) appear to favour the lack or non-obligatoriness of plural marking. This can be seen particularly in Africa, where optionality or absence of plural marking is found particularly in the isolating languages of West Africa.

4 The plural-marking hierarchy

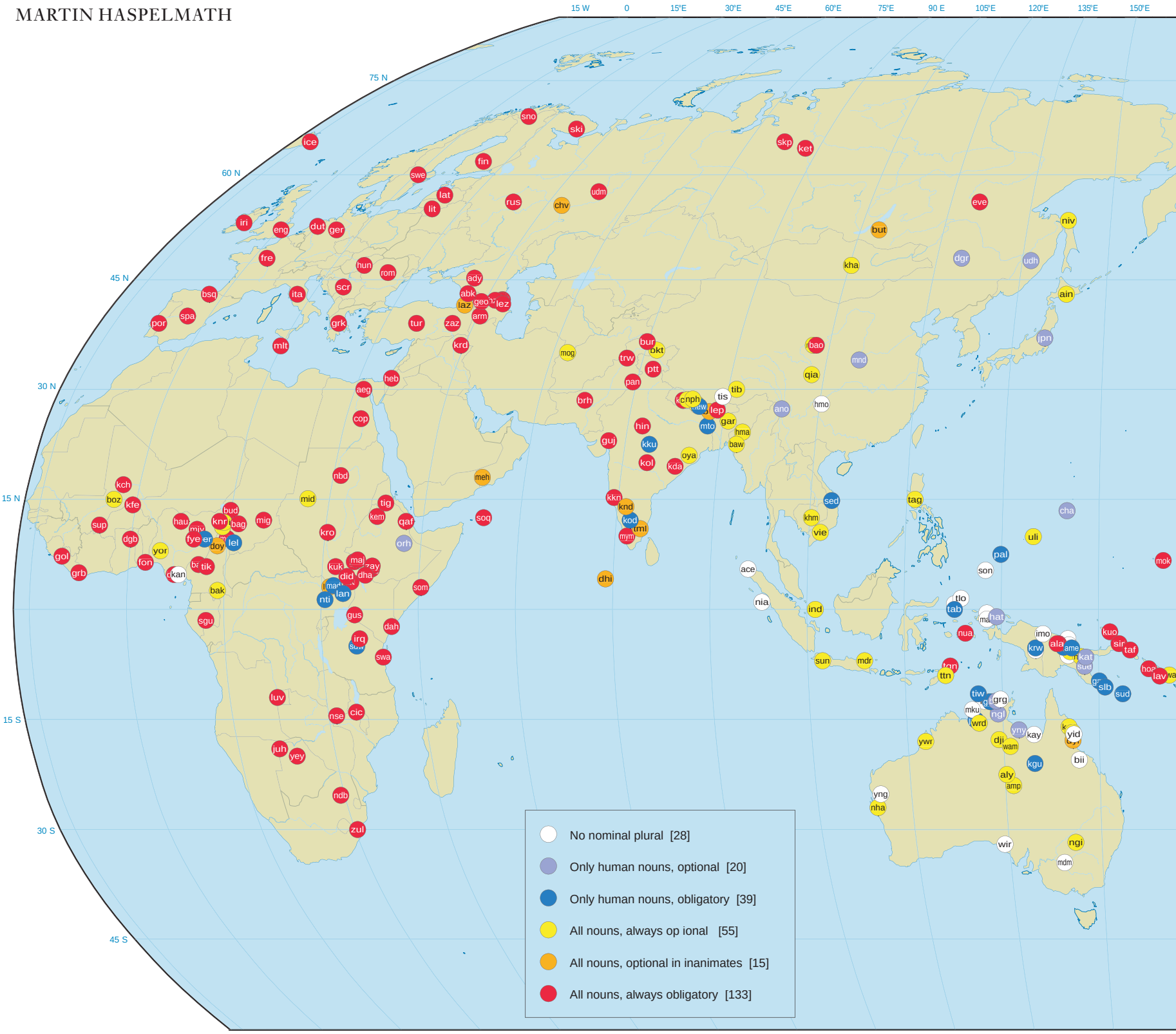
The implicational hierarchy in (1) is part of a larger animacy hierarchy that governs the likelihood of number marking (Smith-Stark 1974, Corbett 2000: §3.1):

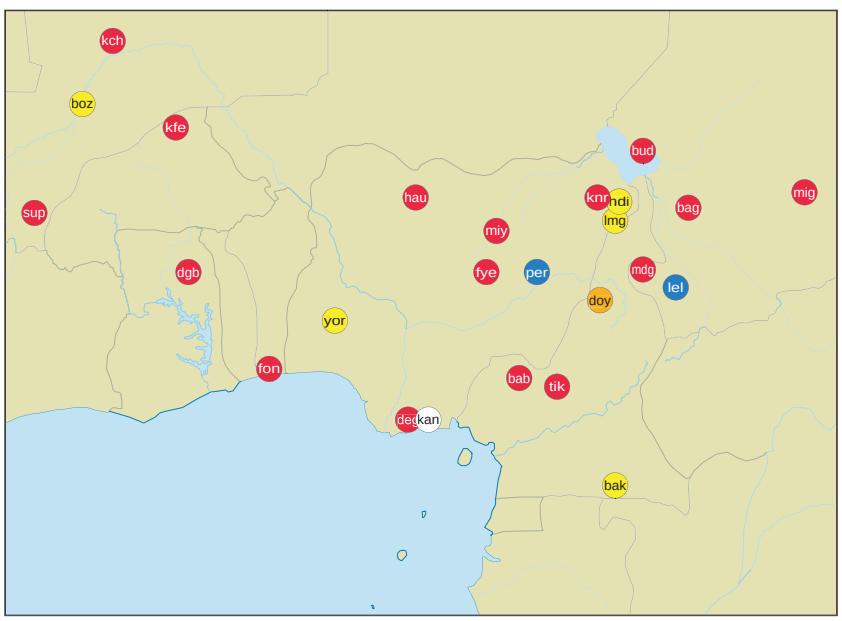
(3) speaker > addressee > third person > kinship terms > other humans > "higher" animals > "lower" animals > discrete inanimates > nondiscrete inanimates

The explanation for the role of animacy in plural marking seems to be the fact that the distinction between one and more than one is more salient for animates than for inanimates, so that speakers are more likely to make use of available plural markers when they refer to a plurality of animates. Through grammaticalization, this preference in language use can lead to obligatoriness in language structure.

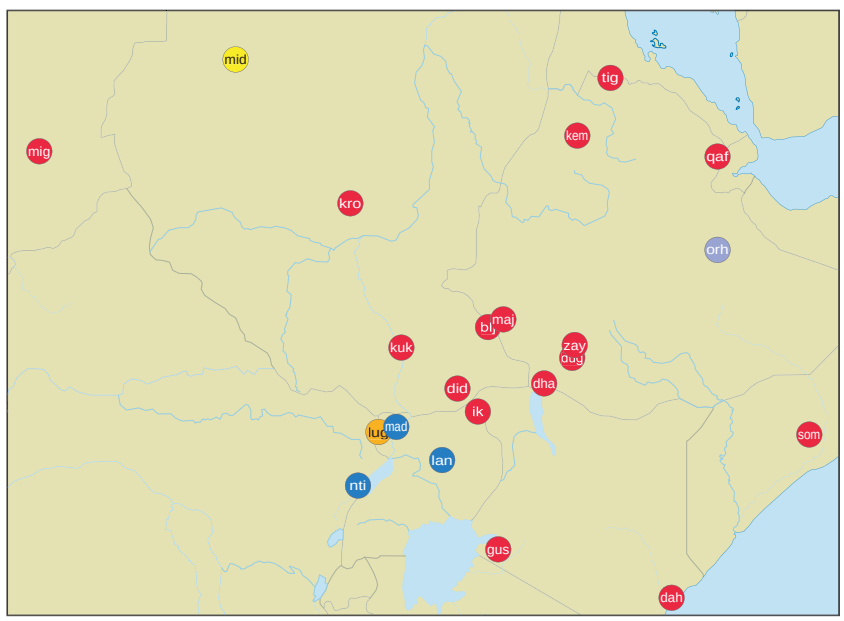
34 Occurrence of Nominal Plurality

MARTIN HASPELMATH

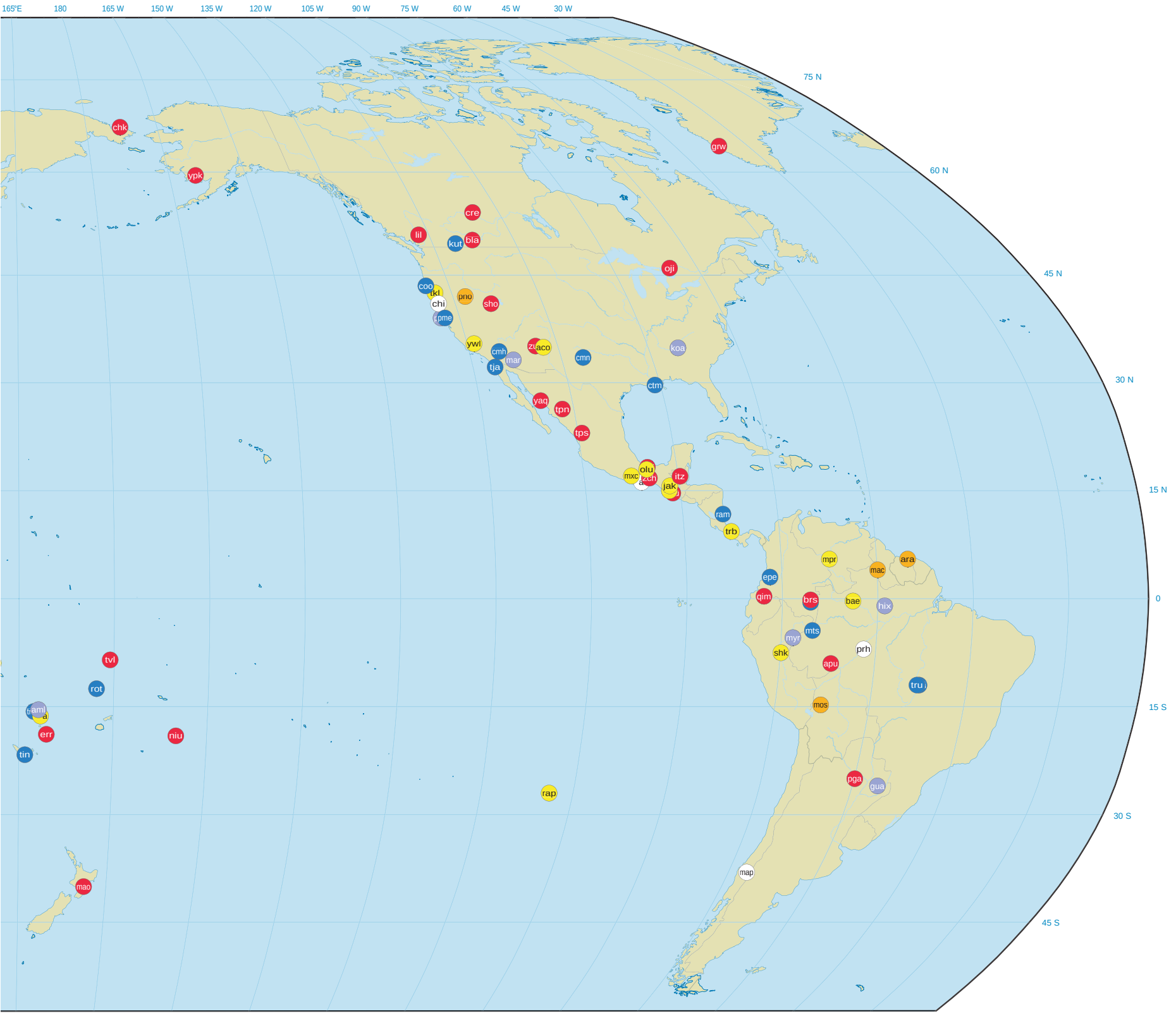


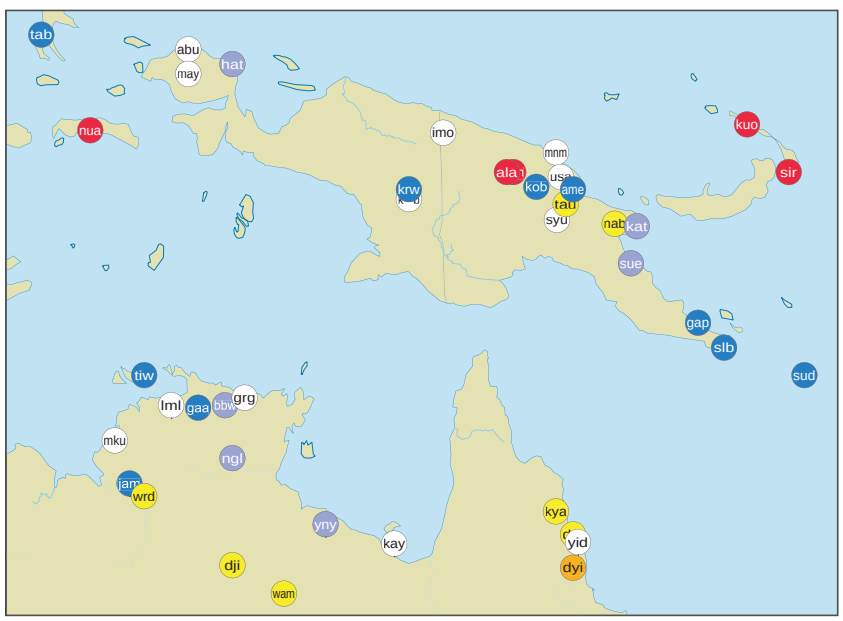


Nigeria and surrounding areas

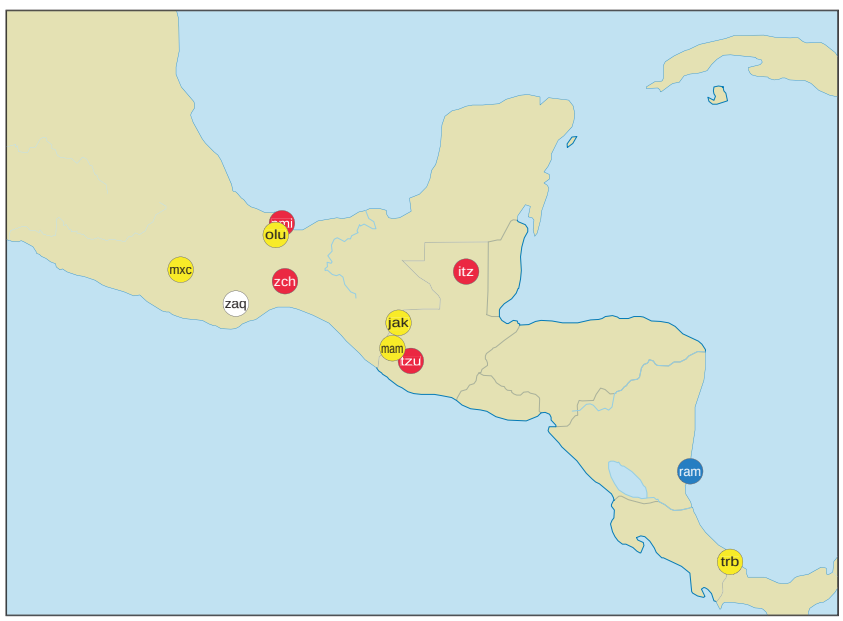


eastern Africa





New Guinea and northern Australia



southern Mexico and Central America

35 Plurality in Independent Personal Pronouns

MICHAEL DANIEL

1 Introduction

Somewhat paradoxically, the plural of personal pronouns has been considered both as a core phenomenon of nominal number and as something that has nothing to do with it. On the one hand, several scholars have suggested that if a language has a number distinction anywhere, then it has it at least in pronouns, or, speaking in diachronic terms, that plurality has spread from plural personal pronouns to other nominals (cf. Forchheimer 1953: 12). On the other hand, linguists have repeatedly argued that pronominal plurals are different from nominal plurals in that the two have a different reference structure. Thus, it is observed that 'we' is not the same as 'I + I + . . . I' (e.g. Benveniste 1966a, Barulin 1980; cf. also Corbett 2000: 83–4). This chapter provides an account of what types of formal correlations between plurals and singulars are actually observed in personal pronouns, and how the devices for expressing pronominal plurality relate to the devices for expressing nominal plurality. It is based on 260 languages, including the 200 languages of the *WALS* sample, and will thus contribute further empirical data to the discussion.

2 Comparability of the data

To ensure cross-linguistic comparability, I singled out one type of pronoun which can be identified in many languages: **independent subject pronouns**. Independent pronouns were chosen because many languages lack non-independent pronouns, and also because they have been most often discussed in comparisons of pronominal plurality with nominal plurality. However, a handful of languages are reported to have no independent personal pronouns, or at least no plural independent personal pronouns. Subject pronouns were chosen for the sake of consistency, because different case forms sometimes display different formal correlation patterns (see §5.2 below).

Both inclusives and third-person pronouns have been excluded from the classification. Third-person pronouns are often deictic rather than personal and only very rarely (in about one tenth of the cases) show a morphological pluralization pattern identical to that of the first- and second-person pronouns (see also Chapter 43). Inclusives are not found in all languages (see Chapters 39 and 40) and often lack number distinctions altogether.

3 Defining the values

1	No independent pronouns	2
2	Number-indifferent pronouns	8
3	Person-number affixes	25
4	Person-number stem	114
5	Person-number stem with a pronominal plural affix	47
6	Person-number stem with a nominal plural affix	22
7	Person stem with a pronominal plural affix	23
8	Person stem with a nominal plural affix	19
	total	260

Type one is represented by two languages reported to have no plural independent subject pronouns or **no independent subject pronouns** at all: Acoma (Keresan; New Mexico) and Wari' (Chapacura-Wanhan; Brazil). Type two consists of languages which use the same form for both plural and singular reference, i.e. have **number-indifferent pronouns**, e.g. Pirahã (Mura; Brazil) *ti* 'I, we'; *gíxai* 'you.SG, you.PL'.

The languages in type three have in common that both **person and number** are expressed by **affixes**, most often cumulatively, so that an affix may not express number without person also being expressed affixally. An example is Mundari (Munda; India) *a-ñ* 'I' ~ *a-liŋ* 'we.DU' ~ *a-le* 'we.PL'; *a-m* 'you.SG' ~ *a-ben* 'you.DU' ~ *a-pe* 'you.PL'. In most of these languages, the stem is common to all persons (like *a-* in Mundari) or conveys only person differences.

The most widespread type, type four, comprises languages combining person and number in an unanalysable **person-number stem,** e.g. Dogon (Niger-Congo; Mali) *mi* 'I' ~ *emme* 'we'; *u* 'you.SG' ~ *e* 'you.PL'. The languages of the fifth type have a **person-number-specific stem** and additionally express plurality **with a pronominal affix** not used on nouns. An example is Amele (Madang; Papua New Guinea) *ija* 'I' ~ *e-le* 'we.DU' ~ *e-ge* 'we.PL'; *hina* 'you.SG' ~ *a-le* 'you.DU' ~ *a-ge* 'you.PL'. The languages in type six use a **person-number-specific stem** and additionally express plurality **with a nominal plural affix**, i.e. an affix also used on at least some nouns (e.g. Russian *ja* 'I' ~ *m-y* 'we'; *ty* 'you.SG' ~ *v-y* 'you.PL'; cf. *slon* 'elephant', PL *slon-y*).

Type seven contains languages which have the same **person stem** in the singular and plural and express plurality **with a pronominal affix** not used on nouns. An example is Chuvash (Turkic; Russia) *epĕ* 'I' ~ *ep-ir* 'we'; *esĕ* 'you.SG' ~ *es-ir* 'you.PL'. Finally, type eight contains languages which have the same **person stem** in the singular and plural **with a nominal affix** to mark plurality (e.g. Mandarin *wǒ* 'I' ~ *wǒmen* 'we'; *nǐ* 'you.SG' ~ *nǐmen* 'you.PL').

4 Explaining the classification

There is a noticeable difference between personal pronouns and nouns with respect to the preferred locus of number expression. Expressing nominal number by changing the nominal stem (as in English *person* ~ *people*) is exceptional (see Chapter 33). Wherever this pluralization strategy exists, it is available only for a small closed subset of nouns. In pronouns, by contrast, the expression of number by a special plural stem is the pattern which dominates in our sample. This provides a first basis of classification of pronominal number systems (reflected on the map by the colour of the symbol)—the distribution of person and number marking between the stem and the affix. One can see that most languages express person in the stem (except some languages of type three), while number may be expressed either affixally (types three, seven and eight), or cumulatively with person in the stem (type four), or both (types five and six). What I did not find is a language where number alone is expressed by the stem without person being also expressed in it. If one assumes that the stem belongs to the domain of lexical meaning while the affix belongs to the domain of grammatical meaning, it may be claimed that in most systems of independent personal pronouns the person is categorized as lexical meaning; pronominal number may be categorized either way.

Person and number expression may be intertwined in such a manner that an obviously regular correlation between singular and plural personal pronouns may hardly be describable segmentally in terms of a stem and an affix (cf. Majtinskaja 1969: 180). This pattern of **nonsegmental correlation** occurs, for instance, in Kolyma Yukaghir (Siberia) (*met* 'I' ~ *mit* 'we'; *tet* 'you.SG' ~ *tit* 'you.PL') and in Ekari (Papua, Indonesia) (*ani* 'I' ~ *inai* 'we'; *aki* 'you.SG' ~ *ikai* 'you.PL'); further cases are Turkish, Evenki, Orok, and Murle. For simplicity, such cases had to be classified together with affixal number marking (type seven).

Expressing pronominal number affixally makes it structurally closer to nominal number. But even in those languages which express number in an affix, the formal identity of the plural marker used on nouns and the marker used on pronouns is rare. This provides a second basis of classification (reflected on the map by the shade of blue of the symbol)—the languages are classified according to whether the number affix is used only on pronouns (type five and seven) or also on at least some other nominals (six and eight).

When we have a formal similarity between plural personal pronouns, it is not always clear whether this alone should be considered sufficient to isolate a plural marker. For instance, in the case of German *wir* 'we' and *ihr* 'you.PL', does *-r* qualify as a plural marker specific to personal pronouns proper? We would need independent evidence establishing the stem of the pronoun (for instance, that the stem is identical to the stem used by the singular or some other nonsingular, e.g. dual, pronoun of the same person) or the plural marker (if it is used with other nominal stems). However, neither of these is available for German (nor for Koyraboro Senni, Fur, Kongo, Grebo, Hindi, Hamtai, Arapesh, Usan, Una); for some languages, there is only indirect evidence (Finnish, Hungarian). Another complication is that in many languages, stems used in plural pronouns are obviously related to, although not identical with, the corresponding singular stems, this formal relatedness being structurally halfway between identical stems and different stems. Finally, a plural pronoun may apparently be composed of the singular stem plus a unique affix (this is especially clear in Yurok, Kutenai, Mapudungun, Makah). To make the data interpretation more robust, these languages are treated as showing no correlation between the singular and plural pronouns (type 4).

5 Mixed types

5.1 Person splits. The classification is complicated by the fact that in many languages the two persons behave differently with respect to the classification established above. An example is English, which uses person-number-specific stems in the first person and a number-indifferent stem in the second person. In fact, about 20 per cent of my sample have some kind of person split. Half of these split cases are languages where one of the two persons (two dozen for the second person against half a dozen for the first) contains a number affix while the other does not. In more than half of the split cases, one of the persons has a person-number-specific stem (the first person in two dozen against one dozen for the second person) while the other has the same stem in both singular and plural. Other cases are very rare. To keep the number of distinct types reasonable, the classification in this chapter always goes with the first person.

5.2 Case paradigm splits. An independent personal pronoun may have different case forms. It is not infrequent that these case forms stand in different relations to each other. Compare the Modern Greek oblique forms *emena* 'I' ~ *emas* 'we', *esena* 'you.SG' ~ *esas* 'you.PL' with the nominative forms *egho* 'I' ~ *emis* 'we', *esi* 'you.SG' ~ *esis* 'you.PL'. Apparently, the number correlation is more regular (even though specific to personal pronouns) in oblique forms. Similar splits appear in many languages, including some Turkic, Uralic, and Mongolic languages. Conversely, the oblique forms are less regular than the subject forms in Tümpisa Shoshone (Uto-Aztecan; California). Finally, in Southeastern Pomo (California) first-person pronouns are more regular in oblique forms, while second-person pronouns are more regular in subject forms. For comparability reasons, the map reflects the correlation displayed in the subject pronouns, and disregards possible paradigm splits.

5.3 Number splits. In some languages, different nonsingular numbers behave in different ways. Thus, in Nenets (Uralic; northern Russia) the pronominal dual suffix is specific to pronouns, while plural pronouns use a nominal plural suffix. Again, for comparability the classification only takes into account plurals in the strict sense, because not all languages have several nonsingular numbers.

5.4 Multiple pronouns. A language may have several pronominal forms in the same person-number slot. The classification of the language may then depend on which forms are chosen. Often one of the forms may be considered more peripheral or less pronominal. Many languages have a distinction of polite versus unmarked in the second singular, and some languages such as Basque apparently have the opposite distinction (unmarked versus familiar). In all such cases, the marked form was disregarded. In some languages, nominals are widely used for personal reference. A nominal may expand to pronominal usage at the cost of losing its nominal meaning. Thus, in many South-East Asian languages kin terms are used as subject pronouns even if the kin relation denoted by the term does not actually hold for the speaker and addressee (see Chapter 45). In some languages, nominals may have personal reference in their literal sense, as when the speaker refers to the addressee or to himself by means of a proper name. These "nominal" pronouns were disregarded.

However, some languages have several non-nominal pronominal forms none of which seems to be unmarked. This happens in some South-East Asian languages, e.g. Burmese, which has a rich set of pronominal forms to code different social dimensions. In these cases, the classification reflects the most regular correlation which is found among the pronominal pairs.

5.5 Optional number marking. As with nouns (see Chapter 34), a plural marker on pronouns may be optional. Such optional markers are often the same as those used on nouns (Bagirmi, Aymara, Chalcatongo Mixtec, Gooniyandi, Wardaman, Uzbek, Turkish, Korean, Ladakhi, Nivkh, Itelmen); only Ju|'hoan has an optional plural marker specific to pronouns. Unmarked forms may be number-indifferent (Aymara, Chalcatongo Mixtec), but in most cases number is also expressed elsewhere—either in the stem (Ju|'hoan) or by an obligatory affix (Nivkh). The map disregards the distinction of optional versus obligatory markers.

5.6 Double number marking. The classification is further complicated by the fact that a pronoun may contain two number affixes. Typically, double-marked pronouns use a nominal plural marker preceded by a marker specific to pronouns (e.g. Wardaman). This order is inverted only in Spanish, cf. *no-s-otros* 'we', *vo-s-otros* 'you.PL', where *-s* may be argued to be nominal plural. Often, the nominal marker is optional (e.g. Turkish, see §5.5). Such systems are classified as types six and eight, i.e. as displaying nominal number marking, although this is a disputable decision, especially in cases like Nivkh, whose plural pronouns consist of a number-indifferent stem, a pronominal number marker, and an optional nominal number marker.

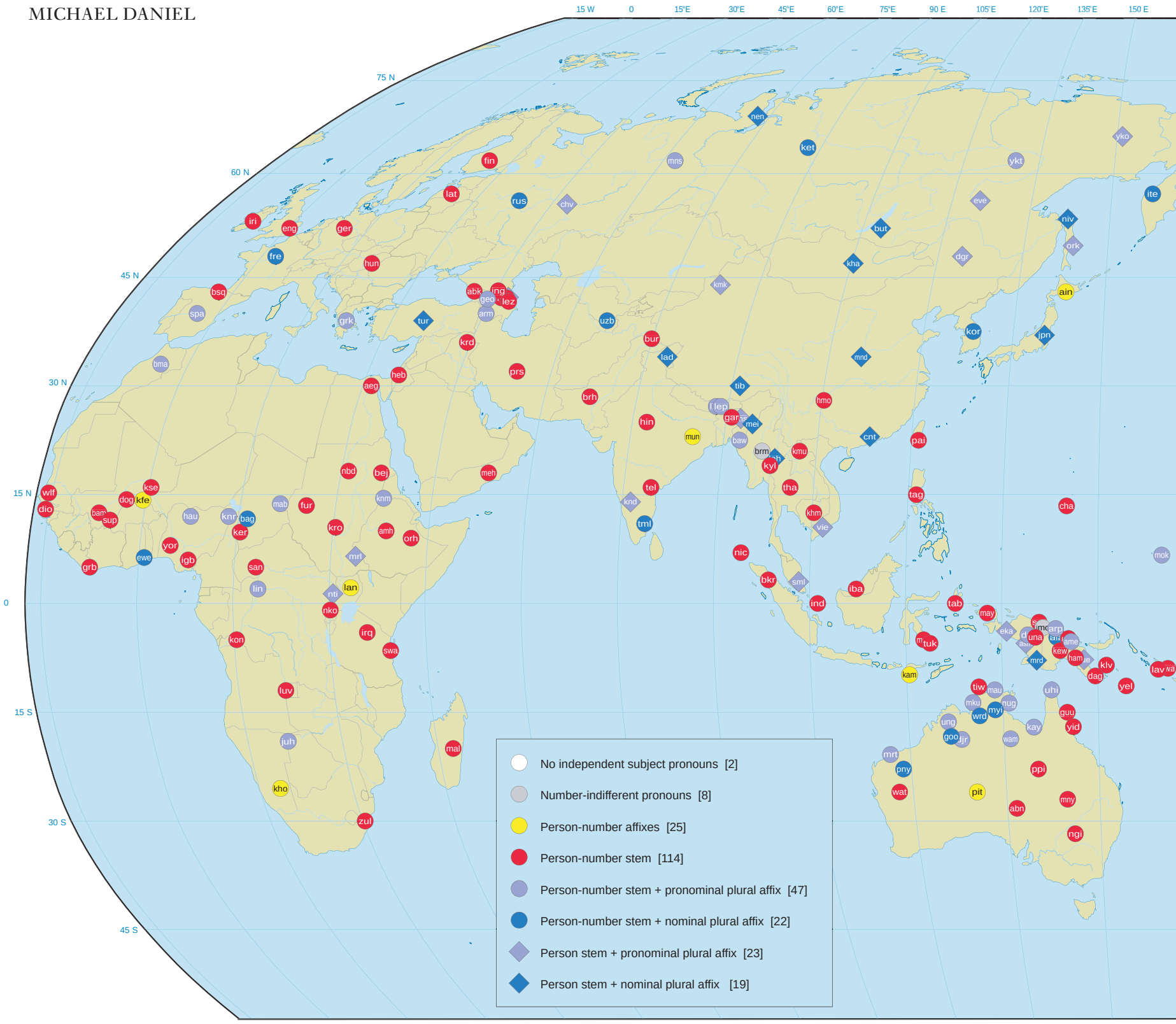
6 Geographical distribution of pronominal systems

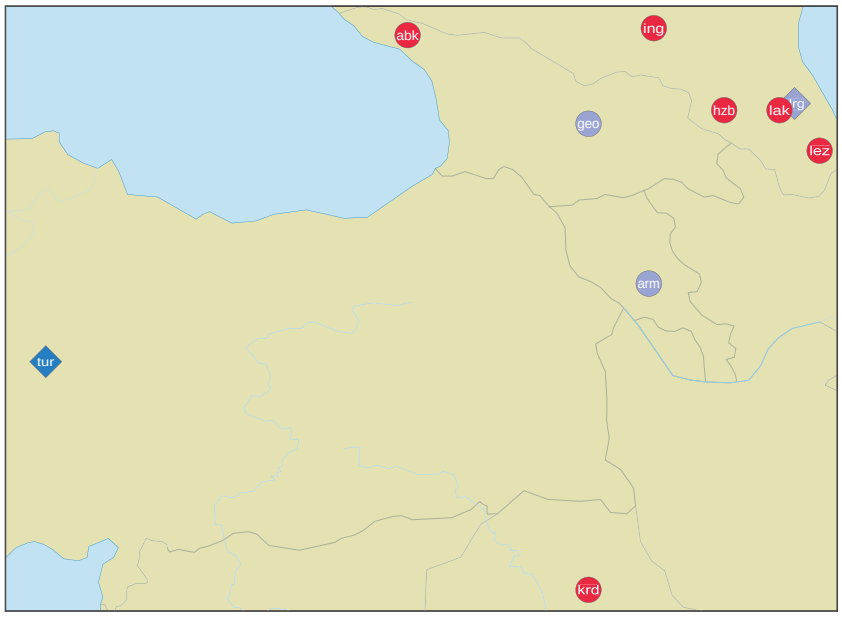
The map shows the following areal patterns. Languages lacking plural independent subject pronouns (type one) are extremely rare; each of the two cases in the sample occurs in one of the Americas (an additional such case, not in the sample, is the African language Mbay, which lacks independent personal pronouns altogether; see Keegan 1997: 62–3). Affixal marking of both number and person (type three) is widespread only in North America; several cases are also attested in Amazonia; only isolated cases occur elsewhere (Ainu, Aleut, Mundari, Koromfe, Lango, Pitjantjatjara, Kambera); some languages show irregular pronominal morphology apparently derivative from this pattern. Pronouns which are indifferent to number reference (type two) are rare; they occur in the Americas and in South-East Asia. Pronouns which express both number and person in the stem without affixal marking (type four) dominate throughout the world, especially in the Middle East, the Pacific, and Africa. Pronouns which use a number affix (types five through eight) are most typical of Asia and northern Australia; they are not infrequent in the Americas and also occur in central Africa. More specifically, pronouns consisting of a person-specific stem plus number marker cluster in eastern Asia and are rare elsewhere, while pronouns consisting of person-number-specific stem plus number marker specific to pronouns are frequent in northern Australia and the eastern Pacific.

Considering areas rather than types, the most homogeneous are Asia (presence of a number affix) and the Pacific (person-number stem with no affix). The eastern Pacific and northern Australia have person-number stems plus number affixes specific to pronouns. In Africa, again, person-number-specifying stems dominate, though central and northern Africa show greater variation. The Americas are diverse in their pronominal patterns; an isolatable number affix is typically absent. In North America there are quite a few languages which have both person and number expressed affixally.

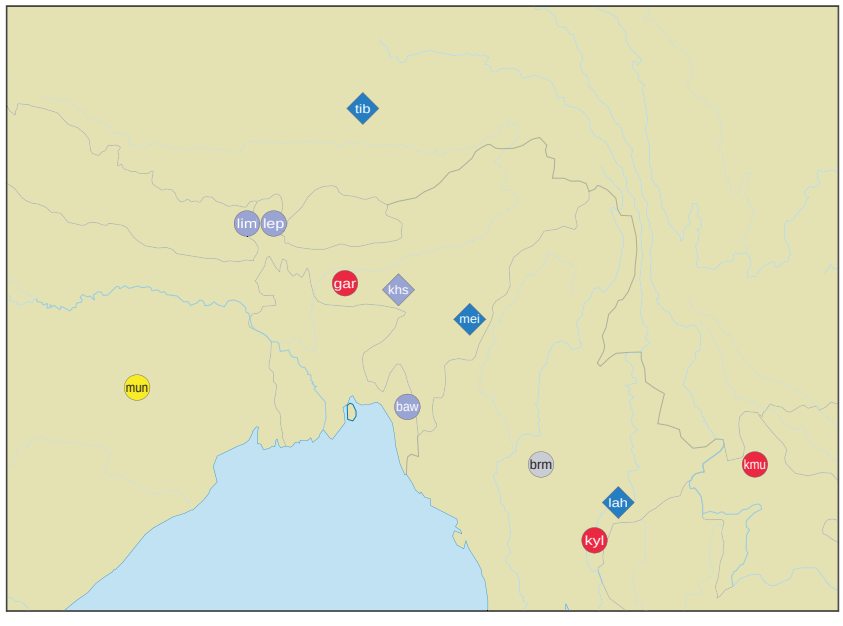
35 Plurality in Independent Personal Pronouns

MICHAEL DANIEL

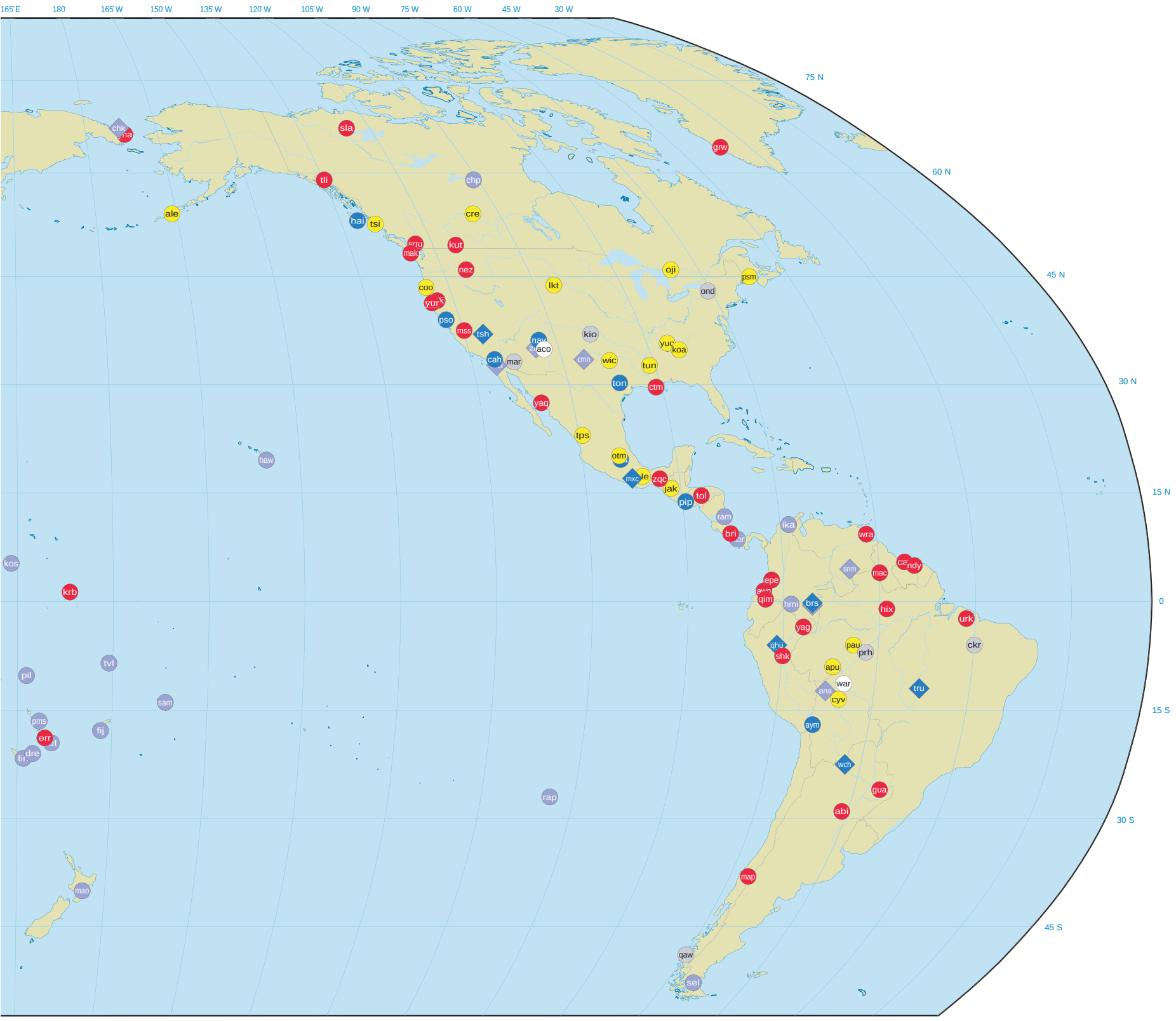




Caucasus and adjacent areas



eastern India and surrounding areas



New Guinea and northern Australia

southern Mexico and Central America

36 The Associative Plural

MICHAEL DANIEL AND EDITH MORAVCSIK

1 Introduction

Associative plural constructions consist of a noun X (typically of human reference, usually a person's name or a kin term) and some other material, most often an affix, a clitic, or a word. The meaning of the construction is 'X and other people associated with X'. An example is Japanese *Tanaka-tachi* 'Tanaka and his associates'.

The reader may have come across various other terms in the literature for the associative plural and related constructions, such as Delbrück's *elliptical dual* (Delbrück 1893: 137), Jespersen's *approximative plural* (Jespersen 1965: 192), the term *plural a potiori* used in Indo-European and Semitic studies, and the label *reprezentativnaja množestvennost'* (representative plural) employed by Russian Orientalists. For extensive discussions of the associative plural construction, see Corbett (2000: 101–11), Daniel (2000), and Moravcsik (2003).

2 Identifying the associative plural

We identify associative plural constructions by two semantic properties: referential heterogeneity and reference to groups. The first property, **referential heterogeneity**, distinguishes associative plurals from additive plurals. **Additive plurals** (cf. Barulin 1980) are the most common type in most languages. An example of the additive plural is English *boys*. It is additive in the sense that it refers to a set where every member is a boy and thus the set is referentially homogeneous: every referent of the plural form is also a referent of the stem. In contrast, the **associative plural** designates a heterogeneous set. For example, Japanese *Tanaka-tachi* does not refer to more than one person named Tanaka: it refers to a group of people only one of whom bears this name. We will call the named member of the group the **focal referent**; the other referents are the **associates**. By virtue of its referential heterogeneity, the associative plural construction is related to other nonhomogeneous plurals, such as what might be called the **similative plural** (e.g. Telugu (Dravidian; India) *puligili* 'tigers and such' (Colin Masica, p.c.)), which differs from the associative plural in that it denotes a class of objects sharing similar features rather than a group of closely related associates. A near-universal instance of heterogeneous plurals is plural personal pronouns, e.g. *we* meaning 'speaker and some other people' (see Chapter 39). The focal referent of associative plurals is often interpreted as pragmatically dominant; thus, in many languages associative plurals meant to designate a family are formed on the husband's (rather than the wife's) name.

The second semantic feature of the associative plural is reference to groups: it refers to a **close-knit group of individuals** rather than to sets without any internal cohesion. This characteristic ties associative plurals to **collective plurals** such as Russian *listva* 'a mass of leaves'; the difference is that collective plurals are referentially homogeneous. The group's numerical size may be two or more. Note that, unless otherwise indicated, we use the term *plural* in a broad sense covering not only plurals but also duals and potentially also trials, paucals, etc., although for associatives we only have examples of plurals and duals.

3 Semantic types of associative plurals

We will consider two aspects of the meaning of associative plurals: the lexical classes that can form associative plurals and the nature of the relationship within the associated group.

3.1 Lexical constraints on associative plurals. What kinds of nouns can form associative plural constructions? In almost all cases, the focal referent and the associated referents belong to the same cognitive category, which is mostly the class of humans or, much more rarely, nonhuman animates (e.g. Tatar (Turkic; Russia) *jeldezler* 'Star (a cow's name) and (my) other cow' (Michael Daniel, field data)). There is only sporadic evidence for inanimate heterogeneous plurals designating sets of objects closely related to each other, and this pattern is never productive. Our examples come from Sanskrit, where the dual form of the word for 'mortar' refers to 'mortar and pestle' (Delbrück 1893: 137), and from Nganasan (Uralic; Russia), where the plural form of the word for 'bow' refers to 'bow and arrows' (Valentin Goussev, p.c.). Among human nouns, associative plurals are much more typical of proper names than of common nouns; and among human common nouns, of kin terms (especially of terms for 'father' and 'mother') more than of other human common nouns.

In some languages, the associative plural construction is not productive: only a closed set of nouns form them. This is the case in Spanish: the few examples include *mis padres*, literally "my fathers" but meaning 'my father and mother', *los reyes*, literally "the kings" but meaning 'the king and the queen'.

All in all, there is a clear preference for associative plurals formed from proper names over kin terms over non-kin human common nouns over nonhuman nouns. Although most of our examples are from the top of this scale, all our evidence is consistent with the generalization that if a language forms associative plurals from any lower type of the four categories mentioned, it also forms associative plurals on any of the higher types (see also Chapter 34).

3.2 Semantic constraints on the associated group. Let us now turn to the second semantic issue: what is the nature of the relation among members of the group that associative plurals refer to? Given X as the focal referent, the group may be: (i) X's family, (ii) X's friends, or familiar associates, or (iii) an occasional group that X is a member of. The kin reading seems to be the core meaning of associative plurals: although there are many languages which allow all three interpretations (e.g. Tatar), there are also languages where neither (ii) nor (iii) is available (e.g. Bagvalal (Daghestanian; eastern Caucasus)).

In sum: there are cross-linguistic preferences regarding the semantic types of associative plurals in both respects, with human proper nouns being the most common stem, and kin forming the most commonly understood associates.

4 Formal types of associative plurals

Table 1 lists the various formal types of associative plurals. The first column identifies the type of marking, the second column expands on the nature of the marking and characterizes the structure of the whole expression, and the third column names an example language. The plus sign indicates co-occurrence without regard to linear order.

The map does not show each of these types separately. We have subsumed them under three more general classes; these will be explained in more detail in the next section.

5 Defining the values

The map shows four feature values. The basic division that we made among languages that do have associative plurals is between those where the marker of the construction also expresses additive plurals and those where it does not. Languages of the former type show up in blue regardless of whether the bifunctional marker is an affix, a clitic, or a free word. Languages with a special associative plural marker are red or orange: red for languages whose marker is an affix and orange for languages where the marking is by means of a free or bound word. Languages with no associative plurals or only

Table 1 Formal types of associative plurals

Marking	Construction	Example
dedicated affix	noun + dedicated affix	Brahui
dedicated article	noun + element which occupies article position and has no other functions	Tagalog
dedicated clitic	noun + clitic with no other functions	Adyghe
basic plural marker	noun + affix, clitic, or a free word, which also serves as additive plural marker on most nouns	Turkish
secondary plural marker	noun + affix, clitic, or a free word, which also serves as additive plural marker on a few nouns	Evenki
possessive affix + plural affix	headless plural possessive	Bulgarian
plural pronoun	noun + plural personal pronoun	Mandarin
conjunction	noun + conjunction without the second coordinand	Basque
plural verb form	singular NP (verbal agreement controller) + plural target verb form	Plains Cree

nonproductive (Spanish) or idiolectal (Russian, English) examples of associative plurals are shown in white.

1. Associative plural marker also used for additive plurals	104
2. Special bound associative plural marker	48
3. Special non-bound associative plural marker	48
4. Associative plural absent	37
total	237

Here are some examples of the four types.

Type one: the marker is **also used to express additive plurals** on at least some nouns. This type includes the fourth and fifth subtype in Table 1.

(1) Zulu (Bantu; South Africa; Lewis 1893: 52–3)
affix:

associative plural:	*o-faku*	'Faku and his company'
	(*au-faku*	'Faku')
additive plural:	*o-kova*	'bananas'
	(*u-kova*	'banana')

(2) Hixkaryana (Carib; Brazil; Derbyshire 1985: 132–3, p.c.)
clitic:

associative plural:	*Waraka* ***komo***	'Waraka and those accompanying him'
additive plural:	*hawana* ***komo***	'visitors'

Type two: the marker is **an affix which is not used for marking additive plurals** on nouns. This type includes dedicated affixes, headless possessives, and conjunctions of the affixal kind in the chart above.

(3) Hungarian (Edith Moravcsik)

associative plural:	*Pál-**ék***	'Paul and company'
additive plural:	*toll-**ak***	'pens'
	toll-ék***	'pens'

Type three: the marker is **not an affix and may not express additive plurals**. This includes dedicated articles, dedicated clitics, plural pronouns, nonaffixal conjunctions and agreement constructions in Table 1.

(4) Old Icelandic (Paul Hopper, p.c.; Heusler 1964: 124–5)

associative plural:	*þeir Gizorr*	'Gizorr and his associates'
additive plural:	**þeir griþingr*	'bulls'

Type four: productive **associative plurals do not exist** in the language, e.g. Russian.

We have chosen the two parameters—morphological versus nonmorphological marking and whether the marker also marks additive plurals—for their inherent interest. First, regarding the latter issue: as the numbers given above show, nearly half of our languages belong to the type where the associative plural marker is also a marker of additive plurals. We wanted to highlight this type because the frequent convergence of two distinct plural meanings on a single form is intriguing. There are two possible ways of interpreting this fact. First, it may be an instance of homonymy—an accidental convergence. Alternatively, it may be an instance of polysemy due to the relatedness of the additive and the associative meanings. We believe the second analysis is the correct one; the argument is as follows. The division between those nouns which, when combined with this marker, are interpreted additively, and those nouns which, when combined with the same marker, are interpreted associatively, is roughly the same in different languages: as noted above, it is primarily human proper names and kin terms whose basic plural forms are interpreted associatively. As was suggested in Daniel (2000), this is so because proper names are linked to a strong pragmatic presumption of their referent's uniqueness, at least within a given speech act. The basic, additive plural interpretation is therefore problematic for such nouns. The meaning of the marker is then adjusted to the semantics of the stem, changing the homogeneous plural reference to the heterogeneous one and yielding associative plural meaning. The same argument is applicable to at least some kin terms.

The other parameter—affixal versus free-form marking—seems to have interesting implications: we assume that the affixes in a given language are historically older than its free forms and thus this difference sheds light on the age of the construction.

6 Geographical distribution

One of the most salient features of the map is the near-absence of associative plurals in western Europe. The only other area having a considerable number of languages without associative plurals is the Americas.

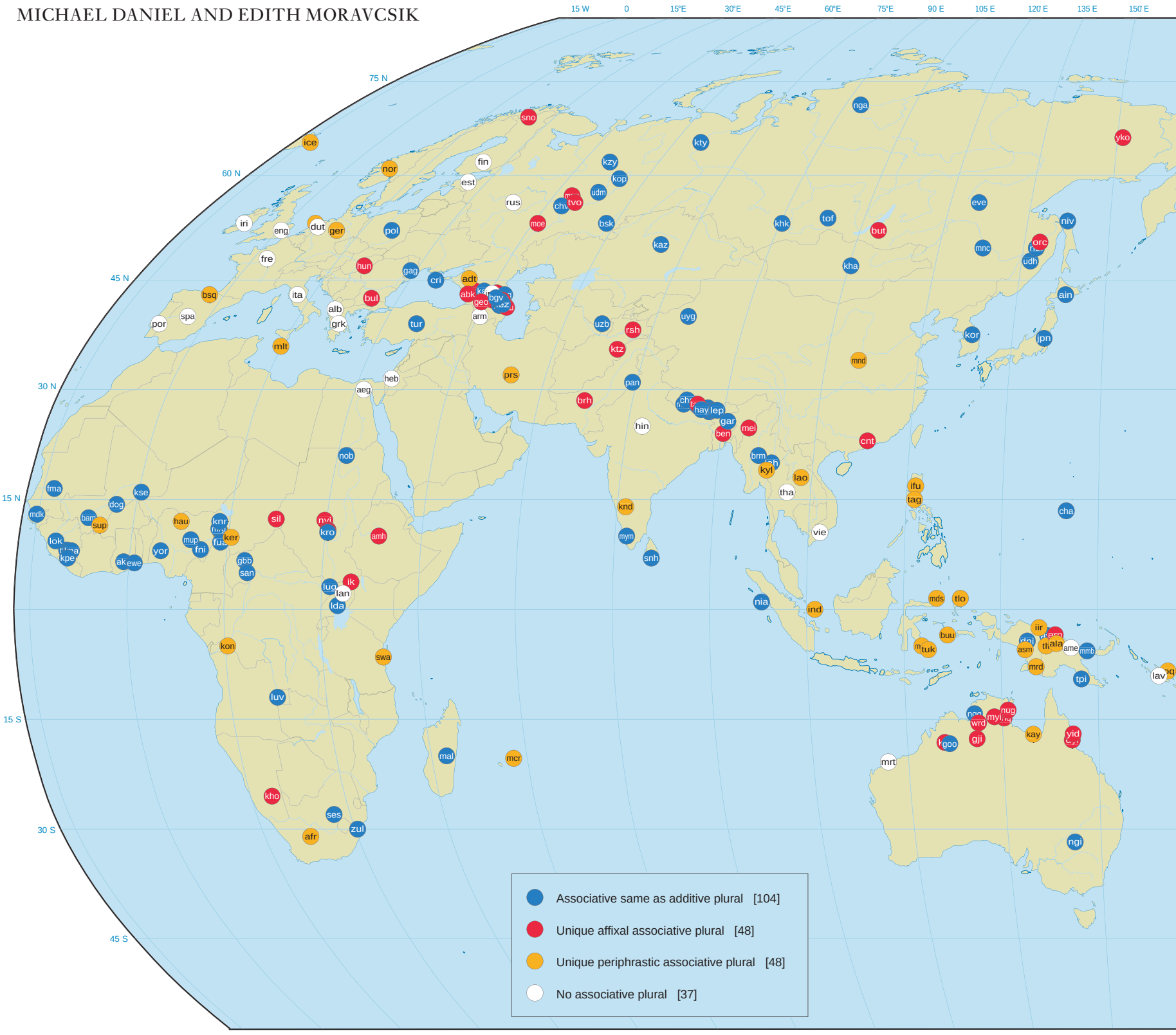
Australia, Asia, and Africa are extremely rich in associative plurals. However, the complete absence of white dots on these continents may be due to the fact that for an underdescribed language where speakers are also not readily available for direct elicitation, it is much harder to confirm the absence of the associative plural than to confirm its existence.

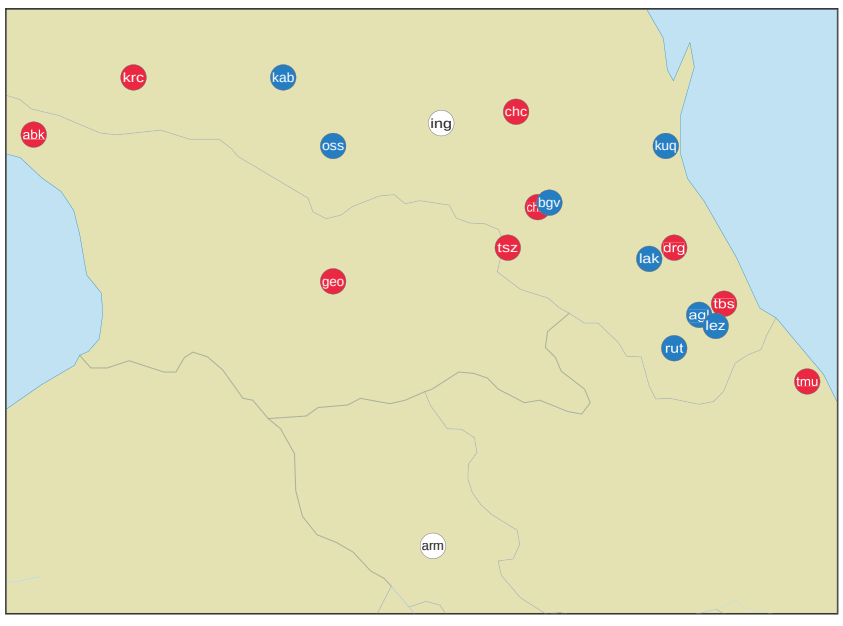
Considering the areal patterning of the formal types of associative plurals, one notices right away that Asia is the area where associative plurals are generally marked by the same means as associative plurals (blue), except for mainland South-East Asia. This type also dominates in Africa and seems somewhat more frequent in the Americas, while it is nearly absent from Australia and the Pacific.

As for associative plurals which are distinct from additive plurals: their bound subtype (dedicated suffix, red) clusters in northern Australia, while the dedicated periphrastic construction (orange dots) is very frequent in the Pacific, which is most likely due to the general scarcity of nominal morphology characteristic of this area.

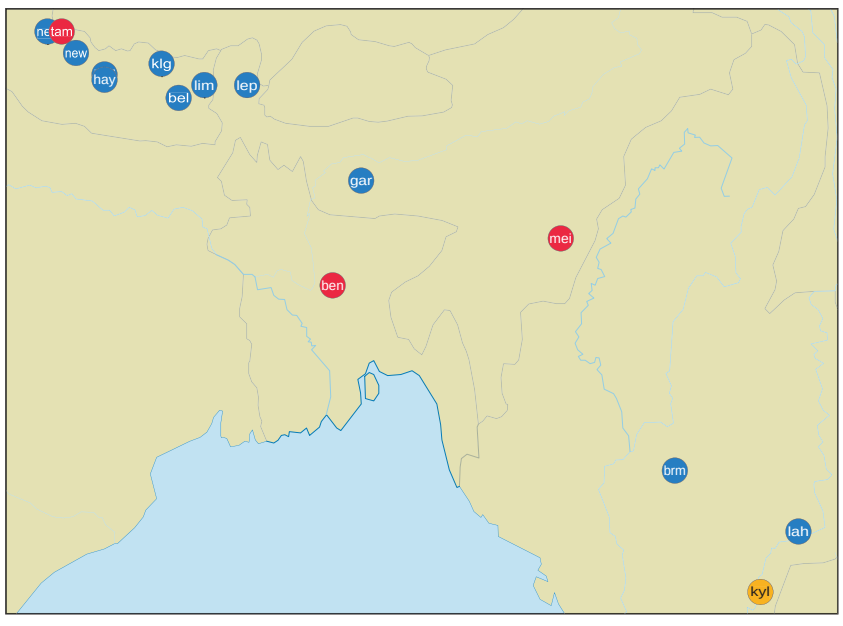
36 The Associative Plural

MICHAEL DANIEL AND EDITH MORAVCSIK

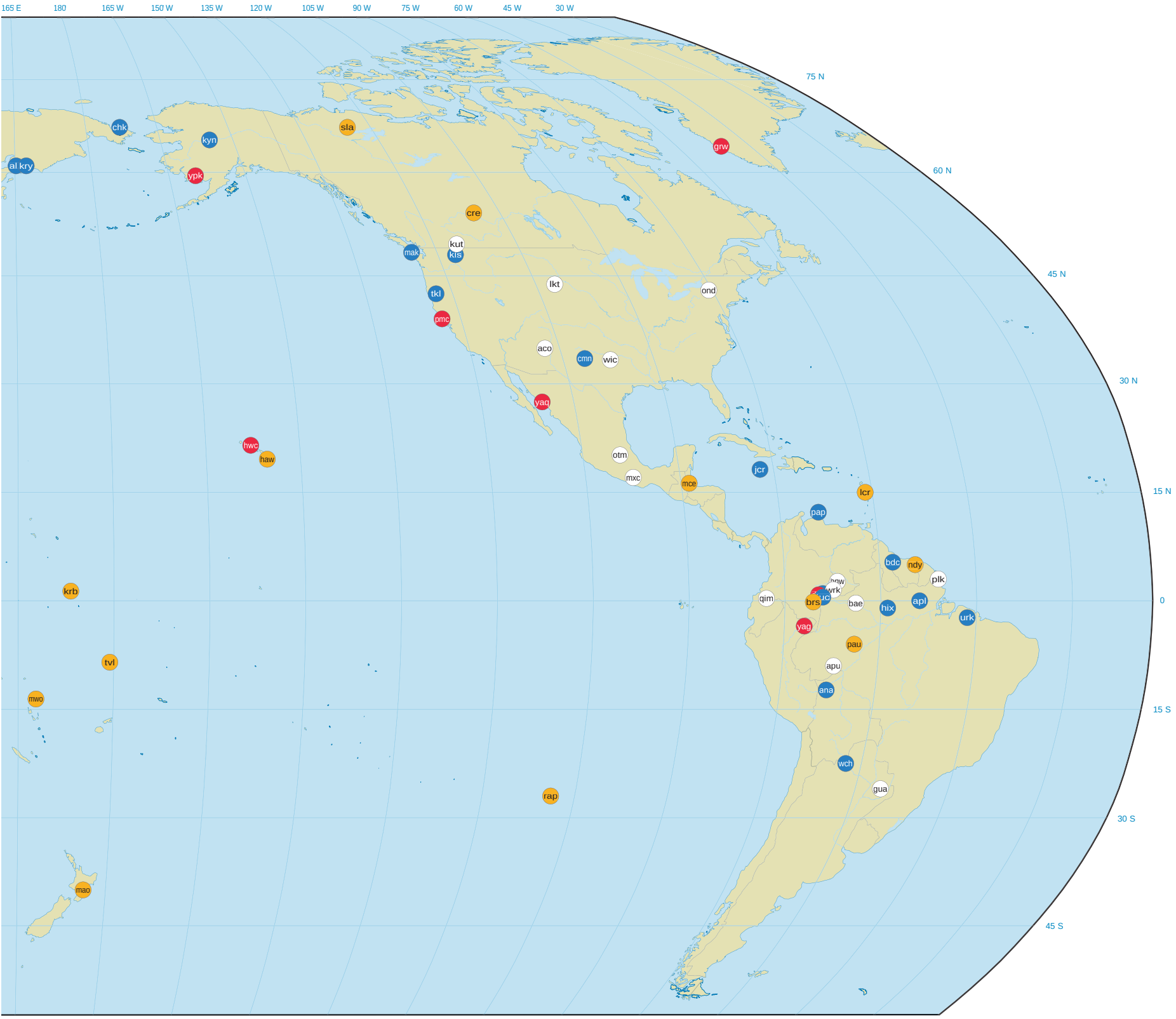


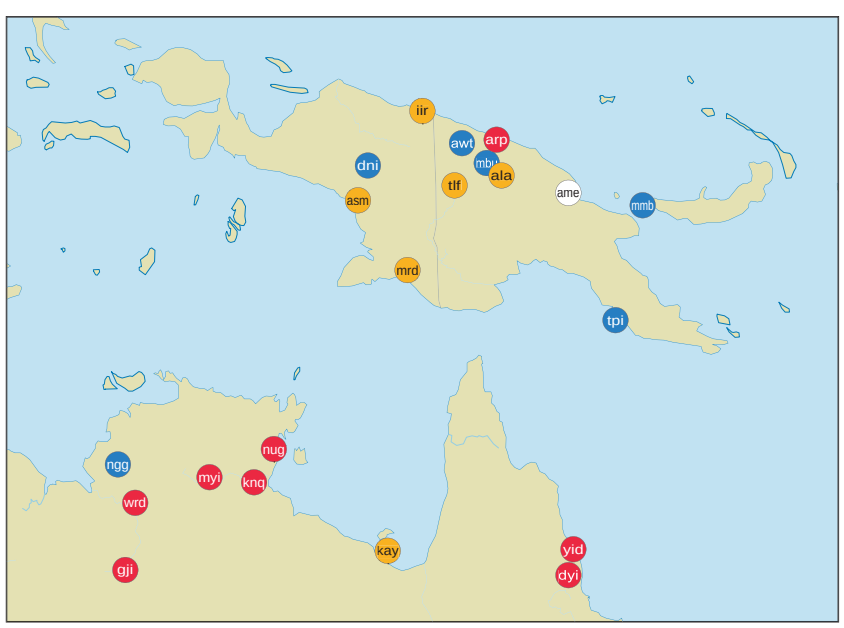


Caucasus

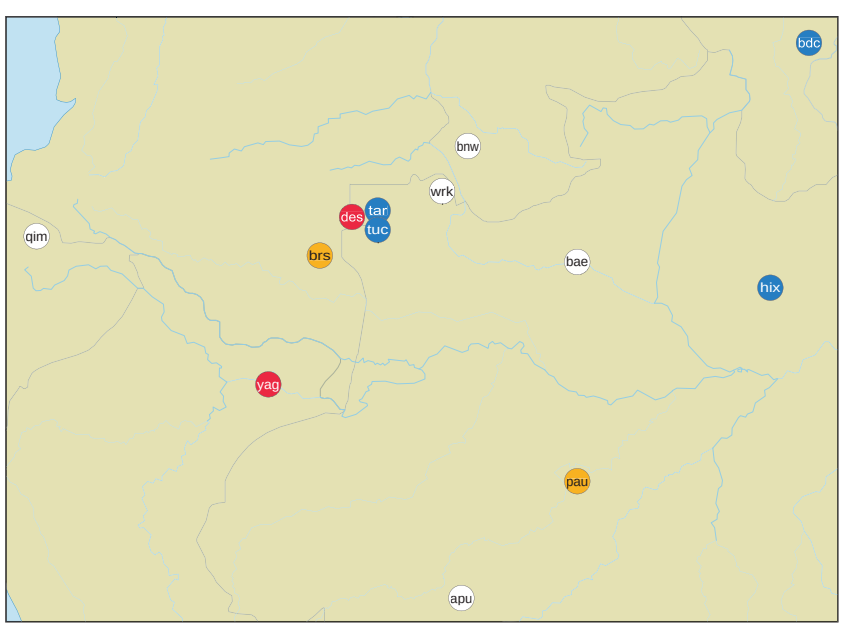


from Nepal to Myanmar





New Guinea and northern Australia



northwestern South America

37 Definite Articles

MATTHEW S. DRYER

1 Defining the values

This map shows information regarding definite articles. For the purposes of this map, a definite article is a morpheme which accompanies nouns and which codes definiteness or specificity, like 'the' in English. This is a somewhat broader use of the term *definite article* than is common, since it includes (i) affixes on nouns that code definiteness; and (ii) demonstratives, if those demonstratives are used as markers of definiteness. In many languages, words that are demonstratives, either in the sense that they can be used deictically with an accompanying gesture or in the sense that they exhibit a distinction in terms of distance (as in the contrast of *this* and *that* in English), are also widely used in contexts where English would use a definite article rather than a demonstrative.

1.	Definite word distinct from demonstrative	197
2.	Demonstrative word used as marker of definiteness	56
3.	Definite affix on noun	84
4.	No definite article but indefinite article	41
5.	Neither definite nor indefinite article	188
	total	566

The first type shown on the map involves languages in which the **definite article is a separate word that is distinct from demonstrative words** in the language. English is an example of such a language: the definite article *the* is distinct from the demonstratives *this* and *that*. Also illustrating this type is Lakhota (Siouan; north-central United States), as in (1).

(1) Lakhota (Ingham 2001: 16)

wic'aša	*ki*	*he*
man	the	that

'that man'

Not only is the definite article distinct from demonstratives in Lakhota, but it can co-occur with a demonstrative, as in (1). Note that languages vary as to whether the definite article precedes or follows the noun.

There are, broadly speaking, two functions associated with definite articles. One of these is an anaphoric function, to refer back to something mentioned in the preceding discourse. The other is a nonanaphoric function, to refer to something not mentioned in the preceding discourse but whose existence is something that the speaker assumes is known to the hearer. This assumed knowledge may be based on general knowledge (as in *the sun*) or it may be based on inferences that the hearer can make in context (for example, inferring from mention of a house that the house has a door, thus making it possible to use a definite article in referring to the door of the house). In some languages, the morphemes treated here as definite articles appear to be restricted to anaphoric usage in that descriptions assign them translations like 'previously mentioned'. An example of such a word is found in Mangarrayi (isolate; Northern Territory, Australia); the word is formed by adding a special prefix (*gi-*) to the distal demonstrative stem, as in (2).

(2) Mangarrayi (Merlan 1982: 44)

ja-guṛwa-n	*gi-nara*	*Ø-gigmuli*
3SG.3SG-circle-PRES	DEF-that	M.ACC-boy

'He goes around the boy.'

Note that in some languages, the definite article is distinct from the demonstrative words but is identical to a third-person pronoun. The examples in (3) illustrate this in Loniu (Austronesian; Admiralty Islands, Papua New Guinea); compare the use of the third-person singular pronoun *iy* in (3a) with its use as a definite article in (3b), where it is distinct from the demonstrative word that follows the noun.

(3) Loniu (Hamel 1994: 76, 100)

a.

iy	*kiʔi*	*la*	*ɛlewɛn*
3SG	PERF	go	far

'He went far.'

b.

iy	*amat*	*iyɔ*
the	man	this

'this man'

The second type shown on the map involves languages in which **one of the demonstrative words is frequently used as a marker of definiteness**. In many languages, it is possible to use demonstrative words anaphorically to refer back to something mentioned in the preceding discourse. However, languages differ considerably in the frequency with which demonstratives are used in this way. In some languages, such usage is relatively unusual, in that anaphoric noun phrases other than pronouns usually do not occur with a demonstrative. In other languages, such usage is very common: in some languages, the majority of anaphoric noun phrases occur with a demonstrative. While there is a continuum of cases between these two extremes, a language is counted here as a language using a demonstrative as a definite article if examination of example sentences or texts reveals that the language uses demonstratives anaphorically more than English does. Most of the languages shown on the map as using demonstratives as definite articles appear to use demonstratives anaphorically considerably more frequently than English. For example, examination of a short text in Ojibwa (Algonquian; Ontario, etc.) in Nichols (1988: 45–7) reveals that of the eleven noun phrases, seven occur with a demonstrative, and two of the remaining four are pragmatically indefinite (an indefinite article is used in the English translation) while two bear possessive affixes ('my car'). Four of the demonstrative occurrences are references to a bear which is one of the two protagonists in the story. When the bear is introduced, it occurs without additional marking (it is one of the two pragmatically indefinite noun phrases referred to above), but all four subsequent references occur with a demonstrative, one of them given in (4), in which the demonstrative *wa* is glossed as 'that'.

(4) Ojibwa (Nichols 1988: 46)

"mii	*maanpii*	*wii-bkeyaanh"*	*kido*	*giiwenh*
but	here	intend-turn.off.1SG	say.3SG	it.is.said

wa	*mko*
that	bear

' "Well, this is where I turn off," the bear said.'

Similarly, the other main protagonist in the text, a man, is mentioned the first time with a verbal expression, but the two subsequent mentions involve a demonstrative. But all of the Ojibwa demonstratives are translated by 'the' in the English translations and the translations would have sounded unnatural if they had been translated with English demonstratives. Definite noun phrases do not always occur with a demonstrative in Ojibwa texts, but they very often do. Descriptions of languages often do not make clear that one of the demonstratives is commonly used for anaphoric reference, so it is likely that some languages shown on the map as lacking definite articles would under closer examination turn out to be languages in which the demonstrative is used as a definite article.

In some languages in which the demonstrative can be used as a definite article, it occurs in a different position within the noun phrase when it is being used as a definite article. For example, in Swahili, the demonstrative follows the noun when used demonstratively but precedes the noun when used as a definite article. The reverse situation obtains in Ute (Uto-Aztecan; Colorado; Givón (in collaboration with the Southern Ute Tribe) 1980: 288–9), Shambala (Bantu; Tanzania; Besha 1993: 28), and Pa'a (Chadic;

Nigeria; Skinner 1979: 61–2). Such languages are all shown on the map as ones in which the demonstrative can be used as a definite article.

The third type shown on the map involves languages in which **the definite marker is an affix on nouns**, as in (5) from Egyptian Arabic.

(5) Egyptian Arabic (Gary and Gamal-Eldin 1982: 59)
ʔiṭ-ṭajjaar-a *gaaja*
the-plane-F.SG come
'The plane is coming.'

In some languages, the definite marker is a clitic which can appear on nouns or on postnominal modifiers, most commonly on the final word in the noun phrase. Such definite clitics are not treated here as definite affixes, but as definite words, falling into one of the first two types. An example of such a language is Angami (Tibeto-Burman; north-east India), illustrated in (6).

(6) Angami (Giridhar 1980: 93)
lêsâdá *kêvī=ù*
book good=the
'the good book'

It is possible that some of the languages shown as having definite affixes actually have definite clitics instead.

The fourth type shown on the map involves **languages which do not have a definite article but do have an indefinite article**, by the criteria discussed in Chapter 38. For example, Tauya (Madang, Trans-New Guinea; Papua New Guinea) has an indefinite article, as in (7a), but no definite article. The indefinite article is not obligatory, however, and a noun phrase without an article may be either definite or indefinite, as illustrated in (7b).

(7) Tauya (MacDonald 1990: 108, 122)
a. *fanu* *ʔafa*
man INDEF
'a man'
b. *nen-ni* *wate* *amoʔo=pe* *ese-i-ʔa*
3PL-ERG house new=BEN want-3PL-IND
'They want a new house.'

The final type shown on the map involves **languages with neither an indefinite article nor a definite article**. In such languages, noun phrases are generally vague with respect to definiteness, as illustrated in (8) for Cherokee (Iroquoian; North Carolina and Tennessee).

(8) Cherokee (Scancarelli 1987: 190)
ki:hli *u:-skala* *achu:ca*
dog 3SG-bite.PUNCT boy
'The/a dog bit the/a boy.'

A morpheme is treated as a definite article for the purposes of this map if it is more generally a specificity marker, i.e. if it marks not only definite noun phrases but specific indefinites. This is the case for the "definite" article in Sango (Samarin 1967: 62–3) and Futuna-Aniwa (Dougherty 1983: 21–3); see examples of the latter in Chapter 38.

There are a variety of morphosyntactic manifestations of definiteness which are not included as definiteness markers on this map. In some languages, pronominal inflection elsewhere in the clause may depend on the definiteness of the coreferring noun phrase, in that inflection will occur only if the noun phrase in question is intended definitely (though the noun phrase itself will generally not bear any marking for definiteness). For example, Ngiyambaa (Pama-Nyungan; New South Wales, Australia) employs absolutive pronominal clitics in second position in the clause, but the third-person clitics are only used if the absolutive noun phrase is intended definitely. Also excluded here are case affixes or adpositions that occur with an object of the verb but only if it is definite, as in Persian (Elwell-Sutton 1941: 24–5). Finally, the map also excludes morphemes in noun phrases which code definiteness but which only occur in the presence of nominal modifiers. For example, the normal way to code definiteness in Danish is by means of a suffix on the noun, as in (9a); if the noun occurs with an adjective, a separate definite article is used, as in (9b).

(9) Danish (Allan et al. 1995: 54)
a. *mand-en*
man-the.COMMON.SG
'the man'
b. *den* *gamle* *mand*
the old man
'the old man'

Because the word in (9b) is used only when a modifier of the noun occurs, it is excluded here, and Danish is coded according to the structure in (9a), as a language with a definite affix.

2 Geographical distribution

Languages with definite articles which are words distinct from demonstratives are common in Europe, especially western Europe, in a wide belt across central Africa from west to east (though not in the south), in New Guinea and the Pacific, and in Mesoamerica, as well as being scattered in other areas. They are relatively infrequent in much of Asia, in South America, and in North America except near the west coast.

Languages in which demonstratives are used as definite articles are widely scattered, but are relatively more common in the United States, and are relatively infrequent in the Americas south of the United States and in Europe.

Languages with definite affixes are relatively common in parts of Europe, both in an area stretching from Scandinavia to northern Russia and in a couple of languages in the Balkan area, in the Middle East, among North-West Caucasian languages, in the same belt across Africa in which definite article words distinct from demonstratives are common, and in western North America.

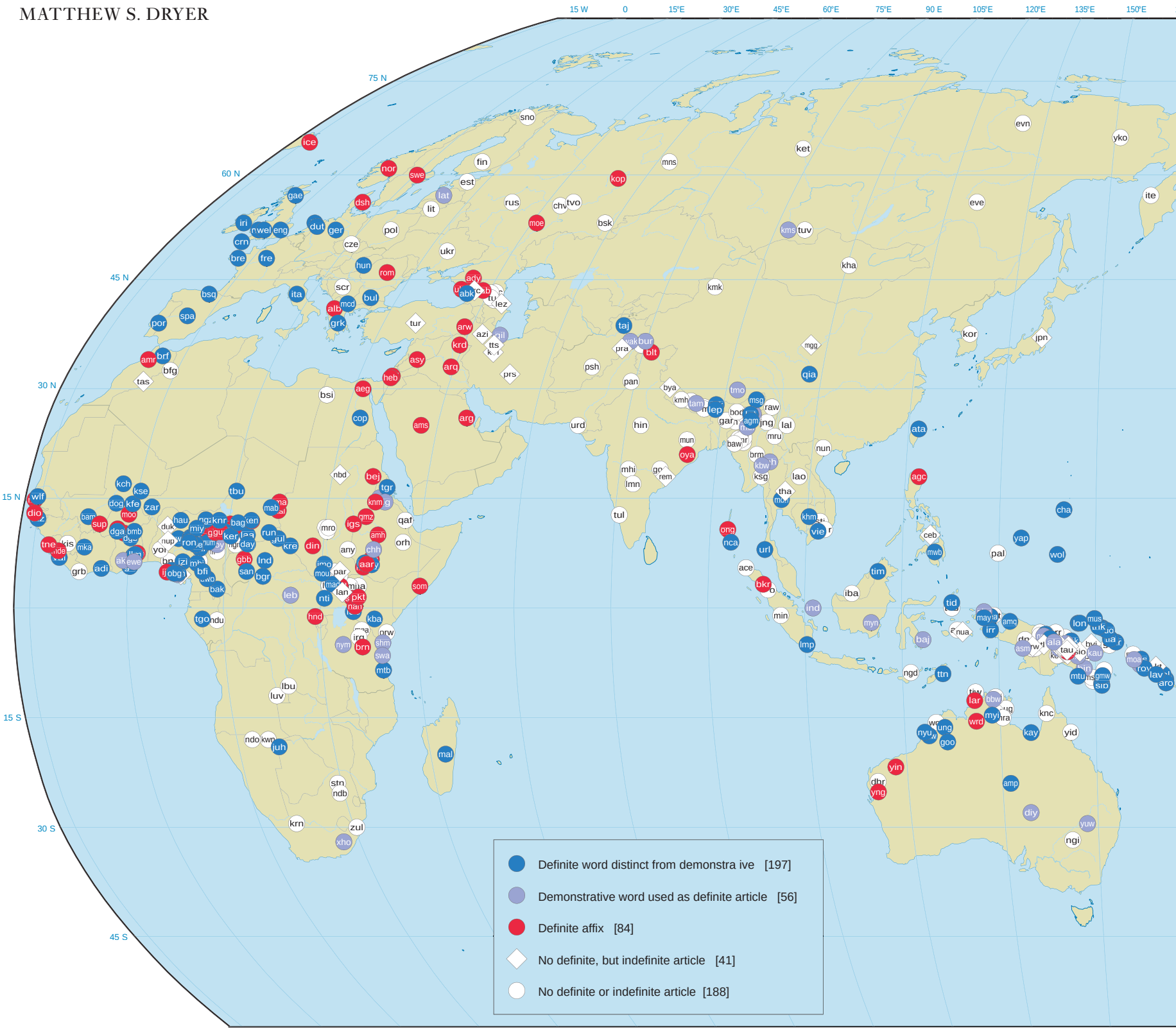
Languages lacking definite articles are common in most areas of the world, but especially so in Asia and South America. One area in which they are infrequent is western Europe. Languages with an indefinite article but no definite article are common in an area in Asia stretching from Turkey to the Caucasus and Iran and in New Guinea.

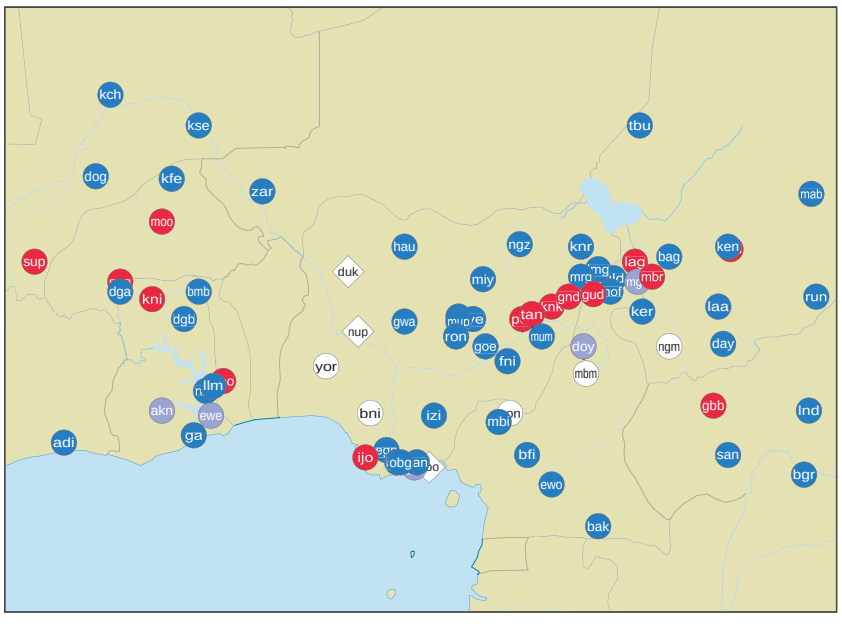
3 Theoretical issues

In English and many other European languages, definite articles are often considered part of a larger category of determiners along with demonstratives and various other words that occur in the same position within noun phrases. In other languages, such a determiner category is less motivated; see Chapters 38 and 88 for discussion. See Map 88 on the order of demonstrative and noun.

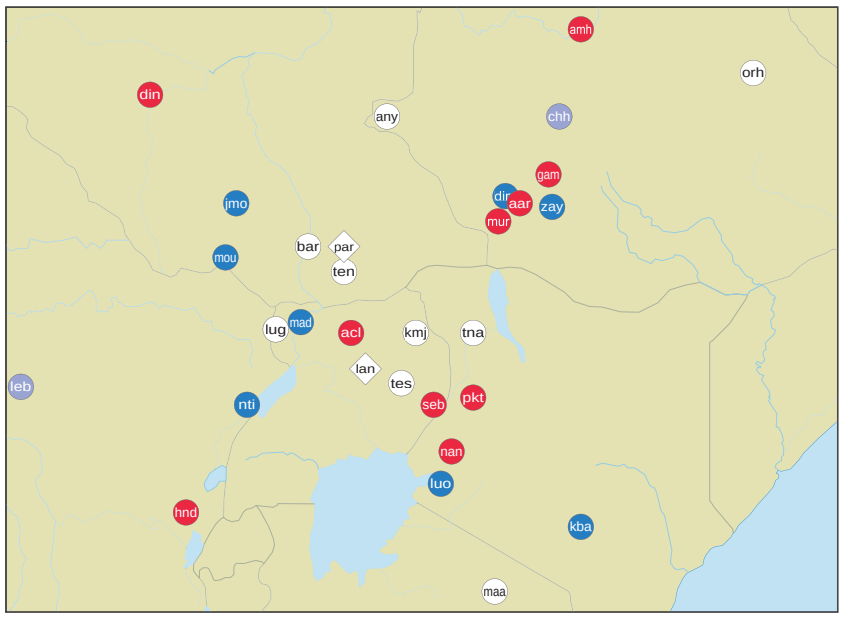
37 Definite Articles

MATTHEW S. DRYER

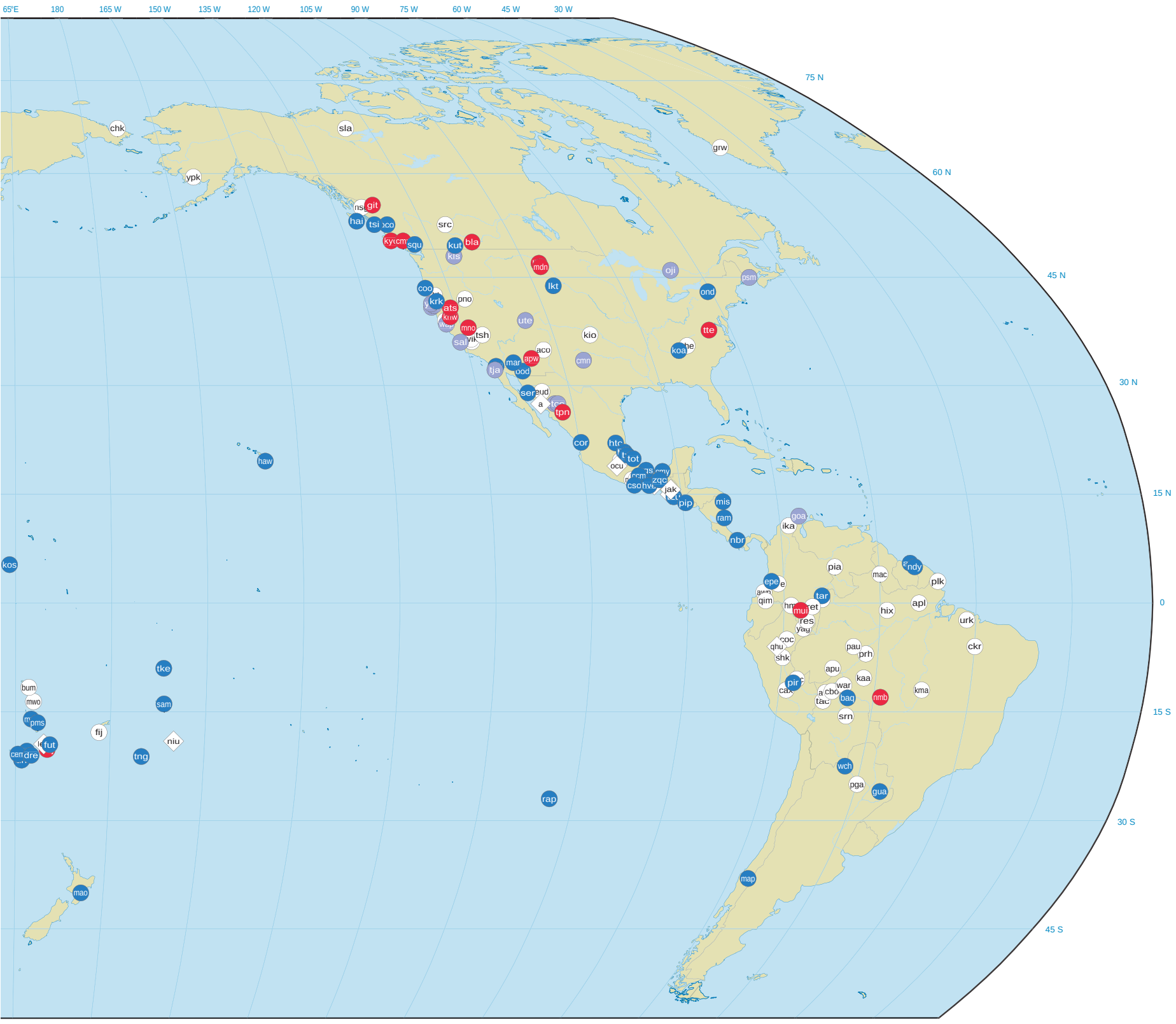


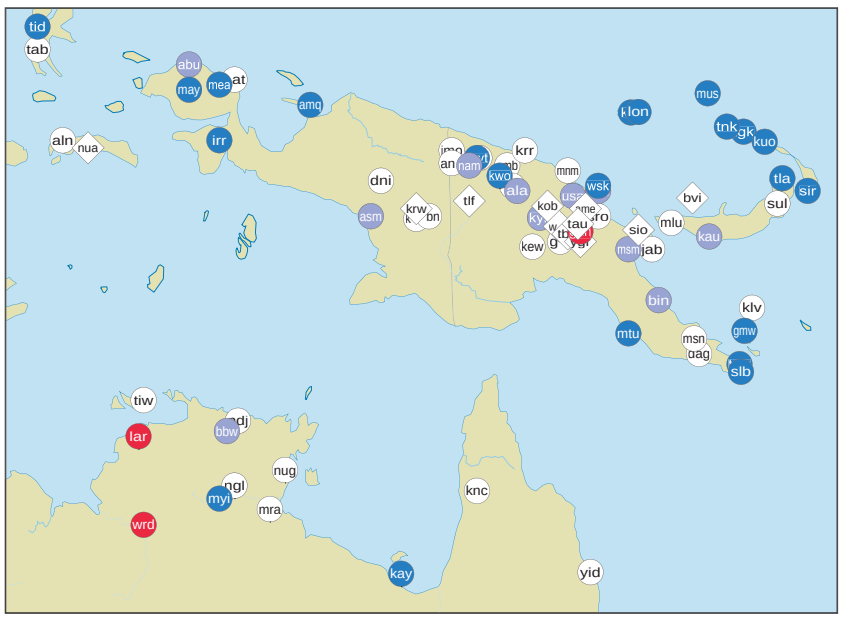


Nigeria and surrounding areas

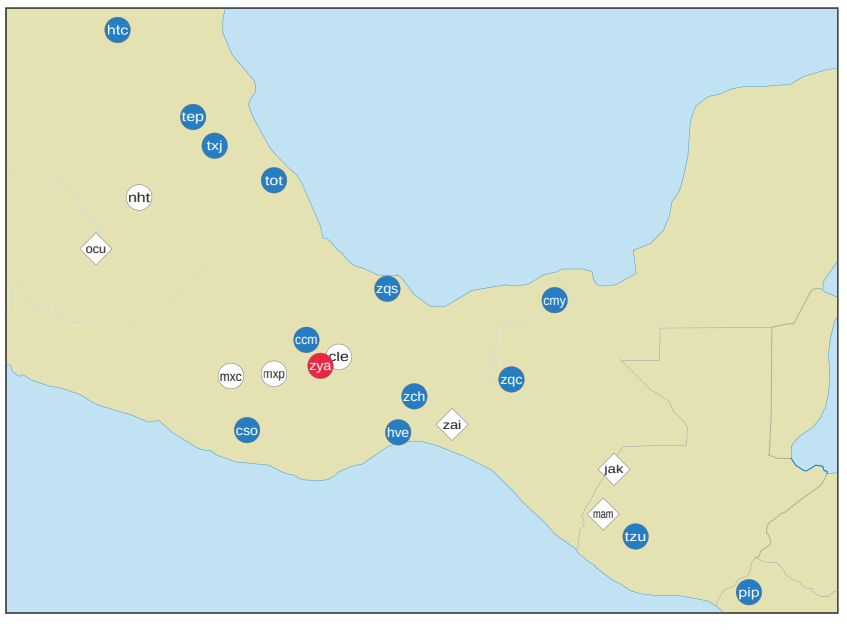


eastern Africa





New Guinea and northern Australia



Mesoamerica

38 Indefinite Articles

MATTHEW S. DRYER

1 Defining the values

This map shows properties of indefinite articles. A morpheme is considered here to be an indefinite article if it accompanies a noun and signals that the noun phrase is pragmatically indefinite in the sense that it denotes something not known to the hearer, like the English word *a* in *a dog*. This includes the use of the numeral for 'one' as an indefinite article and affixes on nouns signalling indefiniteness.

1.	Indefinite word distinct from numeral for 'one'	91
2.	Numeral for 'one' is used as indefinite article	90
3.	Indefinite affix on noun	23
4.	No indefinite article but definite article	81
5.	Neither indefinite nor definite article	188
	total	473

The first type shown on the map includes languages with **an indefinite article that is distinct from the numeral for 'one'**. This includes English, as well as Kobon (Madang, Trans-New Guinea; Papua New Guinea), illustrated in (1); compare the indefinite article *ap* with *añi*, the numeral for 'one' in Kobon.

(1) Kobon (Davies 1981: 61)
nibi *ap*
woman INDEF
'a woman'

The second type shown involves **languages in which the numeral for 'one' is used as an indefinite article**. Many languages use the numeral for 'one' in contexts where English would use an indefinite article. In some languages, this use of the numeral for 'one' is more or less obligatory as a marker of indefinite noun phrases, at least if they are singular. For instance, example (2) from German would be ungrammatical if the numeral for 'one' were left out.

(2) German (own data)
Ich *habe* *einen* *Hund* *gekauft*.
I have one dog bought
'I bought a/one dog.'

In written German, (2) is vague or ambiguous as to whether it corresponds to the English sentence with *an* or the English sentence with *one*. In spoken German, these two senses are distinguished by stress: where English would use *one*, German would stress *ein(en)*. If the only difference in a language is one of stress, then the language is treated as a language in which the numeral for 'one' is used as an indefinite article. In some languages, there are additional phonological differences between the numeral and the indefinite article. For example, in Dutch, the numeral *een* is pronounced with a full vowel [en], while the indefinite article is pronounced with a reduced vowel [ən]. While it is presumably the case that this was originally a single morpheme, which underwent greater phonetic reduction in certain contexts, it is assumed here that this difference has become lexicalized, so Dutch is treated here as a language in which the numeral and indefinite article are distinct. In some languages, the numeral appears in a different position in the noun phrase when it is functioning as an indefinite article. For example, in Turkish, the numeral for 'one' functions as an indefinite article, but has a different position within the noun phrase when serving that function: namely, it follows a prenominal adjective, in contrast to its use as a numeral, where it would precede an adjective (Kornfilt 1997: 275). Similarly, in Remo (Munda; India), the numeral *muy* 'one' precedes the noun when it is functioning as a numeral, but follows when it is functioning as an indefinite article, as in (3).

(3) Remo (Fernandez 1967: 127, 117)
a. *muy* *kaylabay* *gisiŋ*
one black chicken
'one black chicken'
b. *bire* *muy*
stone a
'a stone'

Limbu (Tibeto-Burman; Nepal) is similar to Remo in this respect (van Driem 1987: 32), while the reverse situation obtains in Pa'a (Chadic; Nigeria; Skinner 1979: 55): as a numeral the word for 'one' follows the noun, but as an indefinite article it precedes. Despite the difference in syntactic positions, these languages are coded here as languages in which the numeral for 'one' is used as an indefinite article.

While in languages like German the numeral is more or less obligatory as an indefinite article with singular noun phrases, in other languages the use is clearly optional: examination of two narrative texts in Lezgian (Daghestanian; Caucasus region of Russia and Azerbaijan) in Haspelmath (1993: 445–6) shows that of seventeen instances where the English translations employ an indefinite article in a referential use, only four involve the numeral for 'one' in the Lezgian. Where the numeral as indefinite article is optional in a language, its use often appears to be conditioned at least in part by the discourse prominence of the referent: when something is introduced for the first time and continues to be mentioned in the subsequent discourse, it is more likely to occur with the indefinite article on its first mention, while if something is mentioned only once in a discourse, it is more likely not to occur with the indefinite article (Givón 1981). The same contrast appears to be found in some languages in which the indefinite article is clearly distinct from the numeral for 'one' but still optional.

A language is treated here as a language in which the numeral for 'one' is used as an indefinite article if there is evidence that it is used in at least some situations where English would not use a numeral but would use the indefinite article. For many of the languages shown as using the numeral for 'one' as an indefinite article, this is not mentioned in the description but can be seen in texts. For example, Chelliah (1997) does not appear to mention the use of the numeral as an indefinite article in Meithei (Tibeto-Burman; northeast India), but most of the instances of pragmatically indefinite noun phrases functioning as subject or object in her texts occur with the numeral for 'one'. It is likely that there are other languages like this that are shown on the map as lacking an indefinite article but where examination of texts would show otherwise.

The clearest instances of the numeral for 'one' being used as an indefinite article are in languages in which it can occur as a marker of indefiniteness in plural noun phrases, where the singularity inherent in the original meaning of 'one' is clearly absent. For example, in Lavukaleve (Solomons East Papuan family; Solomon Islands), the numeral for 'one' *ro* not only serves as an indefinite article in singular noun phrases, but also in plural noun phrases, and in the latter event takes plural inflection, as in (4).

(4) Lavukaleve (Terrill 2003: 80)
kanege *rovo*
family INDEF.PL
'some families'

The third type shown on the map consists of languages with **indefinite affixes on nouns**, as in the example in (5) from Korowai (Awju-Dumut, Trans-New Guinea; Papua, Indonesia), where there is an article for specific indefinites, glossed here as 'a certain'.

(5) Korowai (van Enk and de Vries 1997: 75)
uma-té-do *abül-fekha* *khomilo-bo*
tell-3PL.REAL-DS man-INDEF die.3SG.REAL-PERF
'They told that a certain man had died.'

This type does not include languages with indefinite clitics, which may appear on the noun in the absence of nominal modifiers, but which more generally have a fixed position within the noun phrase, in some instances occurring on nominal modifiers. For example, the indefinite morpheme in Dongolese Nubian appears as a suffix on the noun in the absence of postnominal modifiers, as in (6a), but attaches to the postnominal modifier if there is one, as in (6b).

(6) Dongolese Nubian (Armbruster 1960: 323, 327)

a. *hánu=wēr*	b. *kám*	*sɛrɛ́̄=wēg*	*étta*
donkey=INDEF	camel	good=INDEF	bring
'a donkey'	'Bring a good camel.'		

Such indefinite clitics are treated as separate words for the purposes of this map.

The fourth type shown on the map consists of **languages which do not have an indefinite article but do have a definite article**, by the criteria for defining definite articles discussed in Chapter 37. An example of such a language is Kutenai (isolate; western North America). In (7), there is a definite article *niʔ* with the noun *ʔa·kuɬak* 'meat, but it is not obligatory on definite noun phrases: for example, in the absence of a definite article, *yiy̨kiʔmis* 'pot' in (7) can be interpreted as definite or indefinite.

(7) Kutenai (own data)

n̓=uquxaqanmuxu-ni	*yiy̨kiʔmi-s*	*niʔ*	*ʔa·kuɬak*
IND=fall.into-IND	pot-OBV	the	meat

'The pieces of meat fell off into the/a pot.'

The final type shown on the map consists of **languages which have neither an indefinite article nor a definite article**. An example of such a language is Polish: the unmarked noun *jabłko* 'apple' in (8) could be definite or indefinite depending on the context.

(8) Polish (Bielec 1998: 270)

Anna	*je*	*jabłko*.
Anna	eat	apple

'Anna is eating the/an apple.'

As discussed in Chapter 37, the criteria for definite articles include morphemes that signal specificity rather than definiteness: in some languages the morpheme treated in Map 37 as a definite article actually codes specificity, in that it occurs not only with definite noun phrases but also with indefinite specific noun phrases. In the same way, a morpheme that codes a noun phrase as being nonspecific (rather than indefinite) is treated on this map as an indefinite article. For example, the specific article *ta* in Futuna-Aniwa (Polynesian; Vanuatu) is used both where English would use the definite article, and where English would use the indefinite article but where the speaker has a specific referent in mind, as in (9a); by contrast, the nonspecific article *sa* is used where the speaker has no particular referent in mind, as in (9b).

(9) Futuna-Aniwa (Dougherty 1983: 135, 23)

a.

na-n	*tukia*	*ta*	*fatu*
PST-1SG	hit	SPEC	rock

'I hit against a rock.'

b.

a	*roroveka*	*kaseroitia*	*ma*	*sa*	*ika*	*aratu*
ART	Roroveka	catch	NEG	NONSPEC	fish	tomorrow

'Roroveka won't get any fish tomorrow.'

2 Geographical distribution

The features shown on this map do not show strong areal patterns. Languages with indefinite articles of some sort are somewhat more common in (i) Europe, (ii) central and western Africa, (iii) an area in Asia stretching from the Middle East to Myanmar, (iv) the eastern part of New Guinea and the Pacific, and (v) Mexico and Central America. Languages lacking indefinite articles are widely distributed but are somewhat more common in (i) the Americas other than Mexico and Central America, (ii) Australia, and (iii) northern Asia.

Among the three types of indefinite articles shown on the map, languages in which the numeral for 'one' can serve as an indefinite article are widely scattered, while languages in which the indefinite article is distinct from the numeral for 'one' are somewhat more common in Africa and Europe. Languages with indefinite affixes are not common, but are widely scattered. They are more common in Africa than elsewhere in the world.

3 Theoretical issues

As discussed in §3 of Chapter 88, there are some languages, like English, in which articles belong to a category of determiner with a particular position in the noun phrase, while other languages lack such a category in the sense that definite articles and demonstratives occur in different positions in the noun phrase and can co-occur. Similar comments apply to indefinite articles, except that they rarely can co-occur with demonstratives, for semantic reasons. However, there are languages lacking definite articles in which the indefinite article appears in a different syntactic position within the noun phrase from demonstratives, showing that it does not form a category of determiner with demonstratives. For example, the indefinite article in Dongolese Nubian is a clitic that attaches to the last word in the noun phrase, as in (6) above; however, the demonstrative precedes the noun, as in (10).

(10) Dongolese Nubian (Armbruster 1960: 327)

máŋ	*kā́*
that	house

'that house'

In fact, there are languages with both a definite article and an indefinite article in which the two articles appear in different positions in the noun phrase, thus apparently not forming a category in the language. For example, in Jul'hoan (Northern Khoisan; Namibia), the indefinite article follows the noun, while the definite article precedes the noun, as illustrated in (11): in (11a), the indefinite article *n|úí* follows the noun *jù* 'person', while in (11b), the definite article *||'à* precedes the noun *jù* 'person'. (The noun *jù* 'person' in (11b) bears a suffix *-à* 'relative, singular' which is added to nouns when they occur with various sorts of modifiers, including relative clauses and the definite article, though not with the indefinite article.)

(11) Jul'hoan (Dickens 1992: 41)

a.

mí	*hoa*	*jù*	*n|úí*	*ko*	*!aòh*
1SG	see	person	INDEF	OBL	yard

'I saw a (certain) person in the yard.'

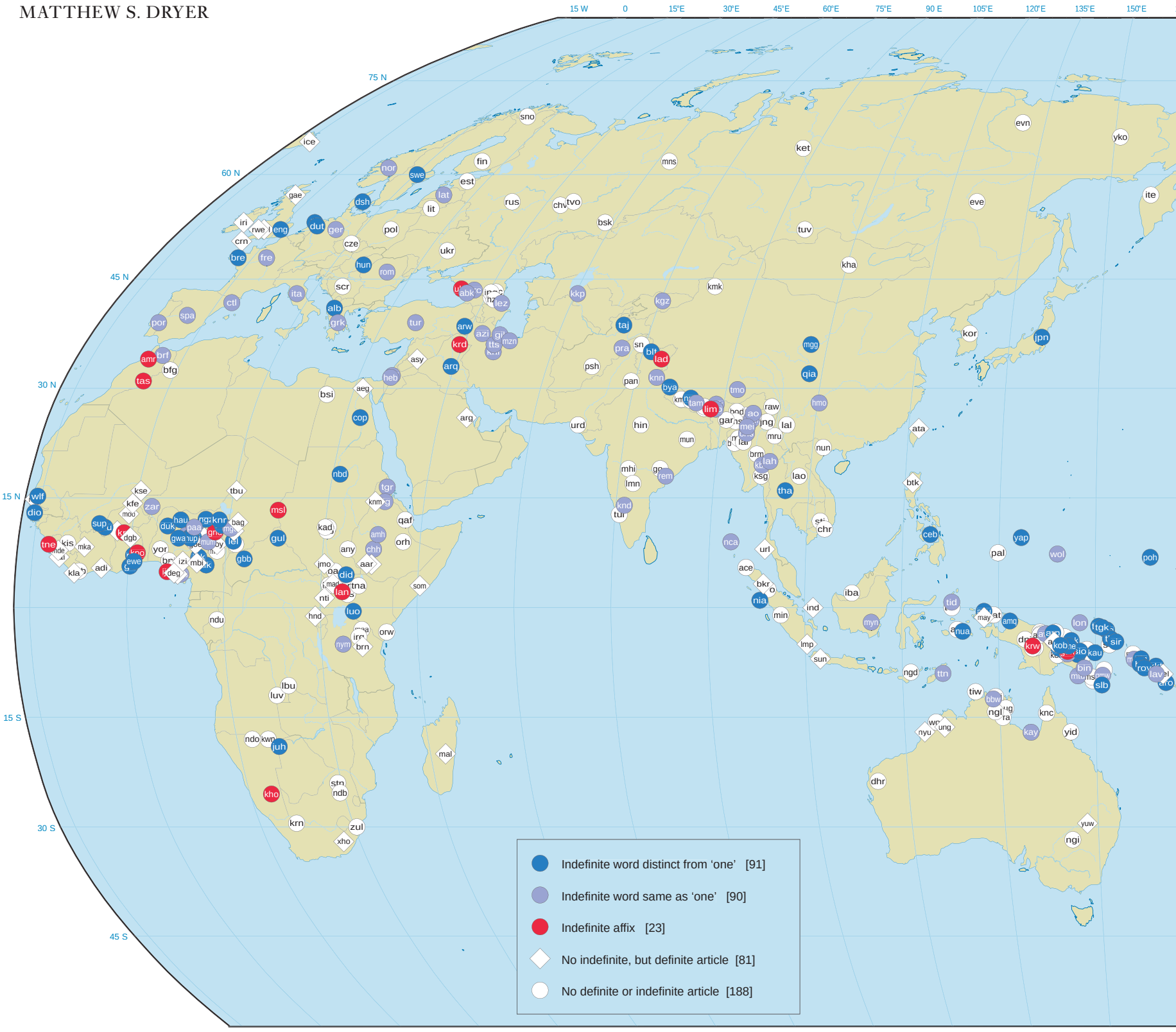
b.

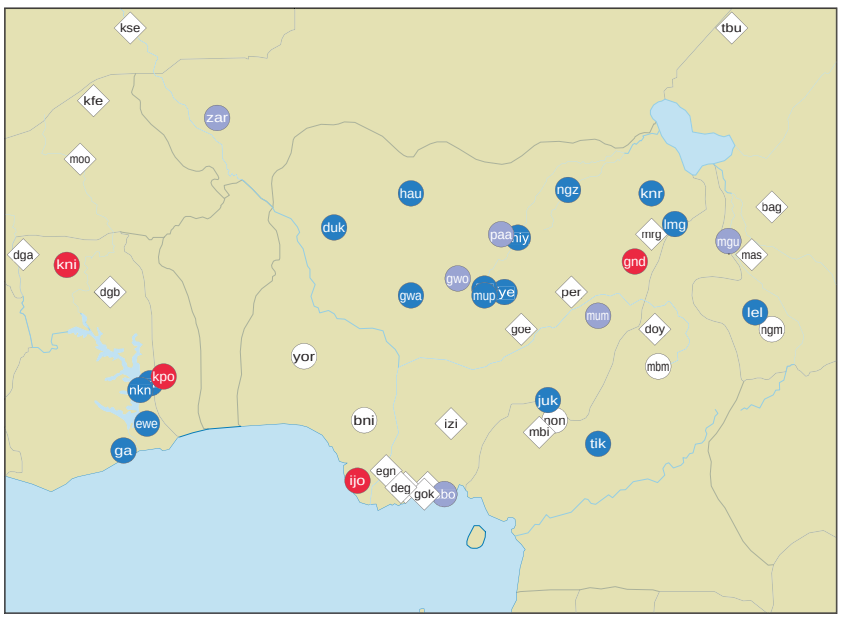
||'à	*jù-à*	*kú*	*!aàh*
DEF	person-REL.SG	IMPF	run

'The person was running.'

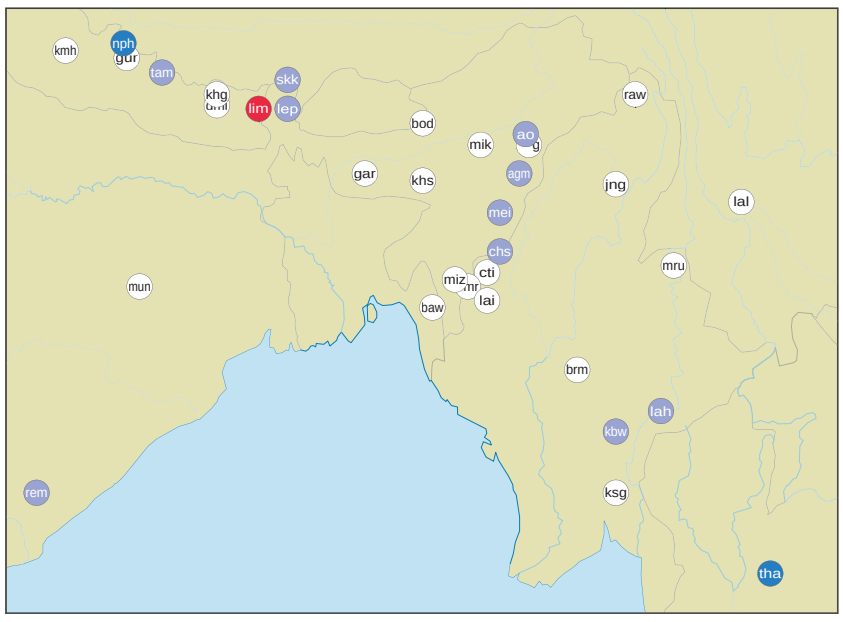
38 Indefinite Articles

MATTHEW S. DRYER

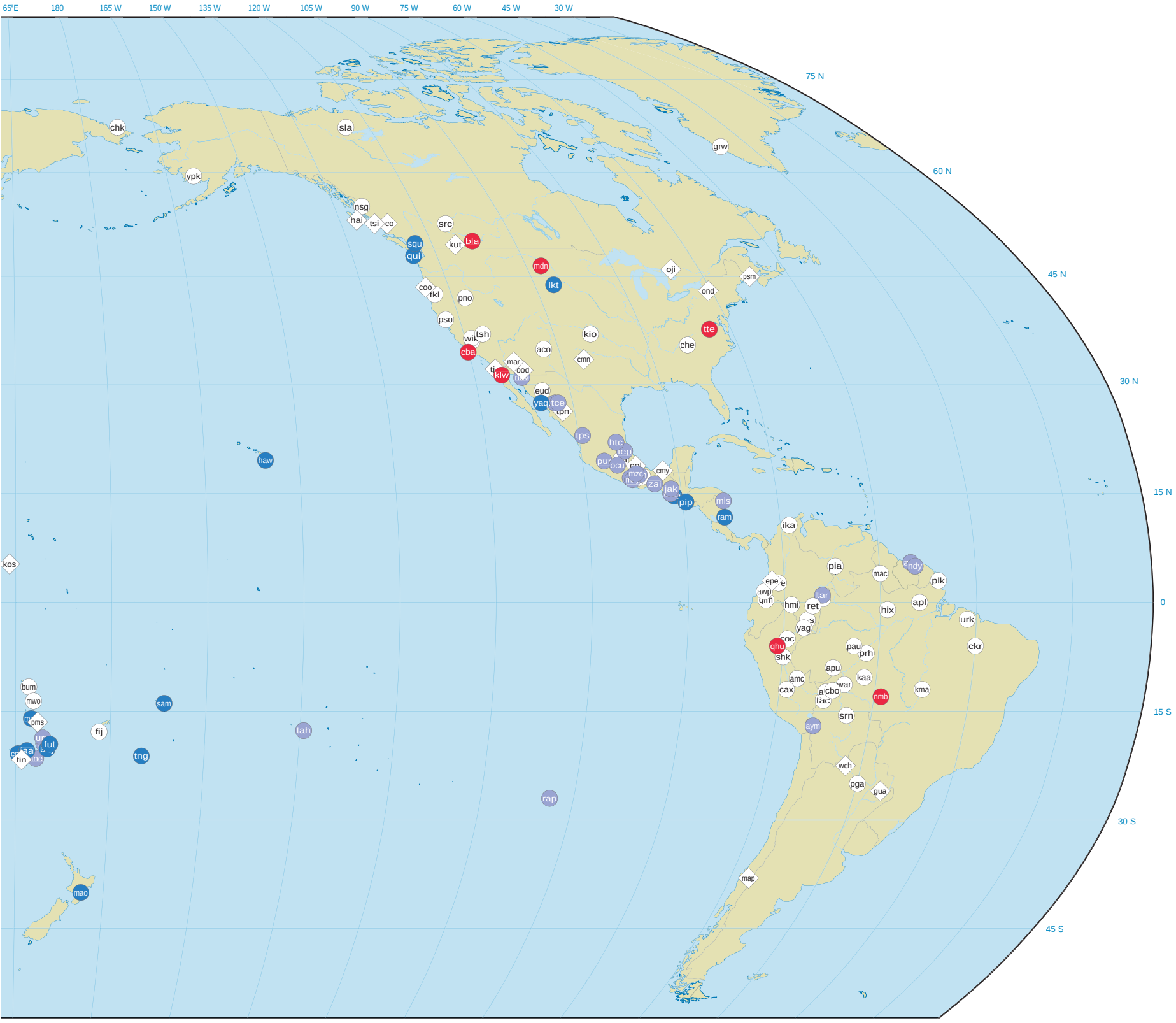


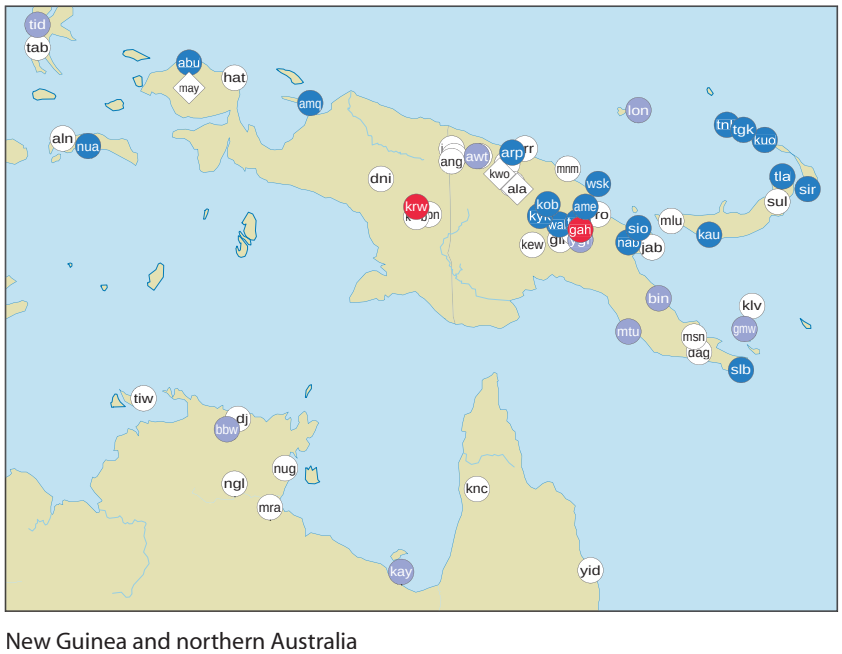


Nigeria and surrounding areas



Nepal, eastern India, and surrounding areas





New Guinea and northern Australia

southern Mexico and Central America

39–40 Inclusive/Exclusive Forms for ‘we’

MICHAEL CYSOUW

General introduction

The distinction between an inclusive and an exclusive pronoun is a commonly attested feature of linguistic structure, yet from a Eurocentric point of view this distinction is particularly exotic. For speakers of English (or any other European language), both the inclusive and the exclusive pronouns are to be translated as *we*. The difference between the two depends on the intended meaning. An inclusive pronoun necessarily includes reference to the addressee. For example, the Mandarin inclusive pronoun *zámen* means ‘we, I and you’; others can optionally be included. An exclusive pronoun, like the Mandarin pronoun *wǒmen*, excludes the addressee from the reference, resulting in a meaning like ‘I and some others, but not you’.

This distinction between inclusive and exclusive is not found in any European language, nor in the languages in its wider surrounding. Because of its absence, this possibility of linguistic structure was not part of any classical linguistic analysis. Its first description dates back to the sixteenth century. The discovery was made by the Dominican friar Domingo de Santo Tomás, as described in his grammar of Quechua, the language of the Incas, first published in 1560. Today, the most widely spoken languages that have this distinction are found among Austronesian languages (in particular Malay and Javanese), among Dravidian languages (in particular Tamil and Telugu), and among northern varieties of Chinese.

39 Inclusive/Exclusive Distinction in Independent Pronouns

1 Definition of values

Map 39 shows the distribution of inclusive/exclusive marking in independent pronouns. The basic distinction is between dots that are coloured red, which mark those languages that have such a distinction, and dots that are coloured blue, which mark those languages that do not have an inclusive/exclusive distinction. However, the actual typology for this map is finer grained: five different types of linguistic structure are distinguished. In the following description of these types, I will use the shorthand ‘we’ for a category that subsumes the inclusive and exclusive meanings (as in the English pronoun *we*).

○	1. No grammaticalized marking at all	2
●	2. ‘We’ and ‘I’ identical	10
●	3. No inclusive/exclusive opposition	120
●	4. Only inclusive differentiated	5
●	5. Inclusive and exclusive differentiated	63
	total	200

1.1 No grammaticalized marking at all. Some languages simply have no basic way of expressing any ‘we’-like concept. Of course, it is possible in all languages to express such meanings, but sometimes there is no specialized means for doing so. For example, in Pirahã (Mura; Amazonas, Brazil) the only way to express a ‘we’-like meaning is by a conjunction of the pronouns for ‘I’ and ‘you’.

(1) Pirahã (Everett 1986: 281)

ti	*gíxai*	*pí-o*	*ahápií*
1.PRON	2.PRON	also-OBL	go

‘You and I will go.’

1.2 ‘We’ and ‘I’ are identical. Some languages have a pronoun for expressing ‘we’, yet this pronoun is the same pronoun as is used for expressing ‘I’. For example, the following sentence from Qawasqar (Alacalufan; Chile) can mean either ‘I ran yesterday’ or ‘we ran yesterday’. There is no way to decide from this sentence alone which meaning is intended. The context must be used to disambiguate this sentence.

(2) Qawasqar (Clairis 1985: 201)

cecaw	*qjeq’ja*	*qjenaq*	*afxat*
1.PRON	run	?	PST

‘We ran yesterday.’ or ‘I ran yesterday.’

In some languages, the pronoun for ‘I’ can be used for ‘we’ (as in Qawasqar), but this usage is uncommon. For example, in Maricopa (Yuman; Arizona) there are no specialized plural pronouns. The existing pronouns “typically [. . .] refer to singular entities. Eliciting overtly plural-marked pronouns is difficult and they appear to be used extremely infrequently” (Gordon 1986: 58). Languages of this type have no inclusive/exclusive distinction, so they are marked blue. However, they are marked light blue because they resemble the previous type with no marking at all (marked white).

1.3 No inclusive/exclusive opposition. This type is well known, as it is the type to which English and many other commonly known languages belong. In this type, there is one pronoun like English *I* and a different pronoun like English *we*, but there is no inclusive/exclusive opposition. Languages of this type are marked in dark blue.

Also included in this type are languages that distinguish a dual without an inclusive/exclusive distinction, for instance Hmong Njua (Hmong-Mien; China and Vietnam; Harriehausen 1990: 124). In Hmong Njua, there are two pronouns for ‘we’, but the distinction is not inclusive/exclusive but dual/plural. The pronoun *wb* is used for dual reference (precisely two persons) and *peb* is used for plural reference (more than two persons).

1.4 Only inclusive differentiated. This type consists of languages that have a special pronoun for the inclusive, but the marking of the exclusive is identical to ‘I’. Such a structure is exemplified in (3) by Canela-Krahô (Ge-Kaingang; Brazil). In (3a), the pronoun *wa* is used in the meaning ‘I’. This same sentence could also be used for the exclusive meaning ‘we (I and some others, but not you)’. However, to express the inclusive meaning ‘we, you included’, a different pronoun *cu* must be used, as shown in (3b).

(3) Canela-Krahô (Popjes and Popjes 1986: 175–6)

a.

wa	*po*	*pupu*
1.PRON	deer	see

‘I see a deer.’

b. *ha* *cu* *jê* *ne* *po* *nõ* *cura*
hey INCL relative and deer ART kill
'Hey, relative, let's go and kill a deer.'

Languages of this type have an inclusive/exclusive distinction which, however, is not fully differentiated lexically, so they are marked pink or light red on the map. There is a clear asymmetry in the structure of the world's languages in that this combination exclusive + 'I' does exist (though it is rare), but that the combination inclusive + 'I' does not exist.

1.5 Inclusive and exclusive are differentiated. The final type distinguished on this map consists of those languages with specialized pronouns for both inclusive and exclusive reference. This is found, for instance, in Chamorro (Austronesian; Guam). The pronoun for 'I' is *hu*, the inclusive pronoun is *ta* and the exclusive pronoun is *in* (Topping 1973: 106–8). This type is marked (dark) red on the map.

There are many languages included in this type that also mark dual number in their pronouns. The basic and most common way to mark duality is exemplified in (4) by the pronouns from Lavukaleve (Solomons East Papuan; Solomon Islands). A special dual pronoun exists both for the inclusive and for the exclusive, both marked by a suffix *-l*.

(4) Lavukaleve (Terrill 2003: 170)

ngai	'I'
el	'exclusive, exactly two'
e	'exclusive, more than two'
mel	'inclusive, exactly two'
me	'inclusive, more than two'

Another strategy is to mark duality only in the inclusive (Plank 1996: 130–1). In such paradigms, the dual inclusive aligns structurally with the singular pronouns, yet strictly speaking it is of course not singular in reference. The term **minimal inclusive** is used to refer to such a dual inclusive. Paradigms with a dual only in the inclusive are known as **minimal-augmented** structures (Thomas 1955). This lexical structure is exemplified in (5) by the pronouns from Southern Sierra Miwok (Penutian; California).

(5) Southern Sierra Miwok (Broadbent 1964: 93)

kan·i	'I'
mah·i	'exclusive, two or more'
ʔoti·me	'inclusive, exactly two'
ʔotic·i	'inclusive, more than two'

The opposite distribution of duality – dual in the exclusive but not in the inclusive – exists among the world's languages, but it is extremely rare (Cysouw 2003: 221–2). In the present sample it is attested in the pronouns from Yagua (Peba-Yaguan; Peru; Payne and Payne 1990: 369–70). Just as exotic is the division attested in Gooniyandi (Bunaban; Australia). Here the inclusive dual is expressed by the same pronoun as the exclusive, *ngidi*, but is different from the inclusive plural *yaadi* (McGregor 1990: 167–73). There are more cases like this among the world's languages, but not many (Cysouw 2003: 93).

Finally, some languages with trial ('exactly three') or paucal ('a few') marking in the inclusive and the exclusive are also included in this type. Trial or paucal marking only occurs among Austronesian languages, but within this linguistic stock it is rather widespread. For example, in the present sample it is found in Paamese (Oceanic; Vanuatu; Crowley 1982: 80).

A special variant of trial marking are pronouns that only distinguish a trial in the inclusive, but not in the exclusive. This is analysed as an extension of the minimal-augmented pattern in (5), adding a category in between 'minimal' and 'augmented' known as **unit-augmented** (McKay 1978). This structure is attested almost exclusively in northern Australia (cf. Cysouw 2003: 232–6). In the present sample this type is attested in Mangarrayi (Merlan 1982: 102).

2 Geographical distribution

In general, the inclusive/exclusive distinction is rather uncommon in Africa and Eurasia. There is no distinction in any language in Europe and its wider surroundings. The nearest cases are a few languages in the Caucasus. The distinction is also relatively uncommon in Africa. In Asia, the Dravidian and the Munda languages have an inclusive/exclusive distinction, although the Dravidian language Kannada in the present sample has lost the distinction under the influence of the neighbouring Indo-Aryan languages. In north-east Asia there might be an area with an inclusive/exclusive distinction, as exemplified here by Evenki, Ainu, and Nivkh. Originally part of this cluster, Khalkha (Mongolian) has lost the distinction, while the northern dialects of Chinese have acquired the distinction. There is an interesting predominance of languages that do not differentiate 'I' from 'we' in South-East Asia. It is rather common in this area for languages not to mark number at all, or to mark it only optionally (cf. Chapter 34).

Off the Asian mainland, the inclusive/exclusive distinction is regularly attested. This is mainly due to the Austronesian languages and the non-Pama-Nyungan languages of northern Australia. In both groups, almost all languages have the distinction. In contrast, it is rather uncommon among the non-Austronesian ("Papuan") languages of New Guinea. Among the Pama-Nyungan languages of Australia the distinction is roughly evenly divided (see next section).

In the Americas, there are about as many languages with an inclusive/exclusive distinction as without it. There seems to be no clear areal division between the two types here. Interestingly, the minor types (lightly coloured) are relatively common throughout the Americas.

3 Pama-Nyungan: a showcase for areality

There are no clear worldwide patterns in the distribution of the inclusive/exclusive opposition. The few patterns that are attested appear to be on a smaller scale. To exemplify the smaller scale, I carried out an in-depth investigation of the areal distribution of the inclusive/exclusive opposition among the Pama-Nyungan languages from Australia, as presented in Map 39A.

1. No inclusive/exclusive opposition	31
2. Inclusive and exclusive differentiated	40
total	71

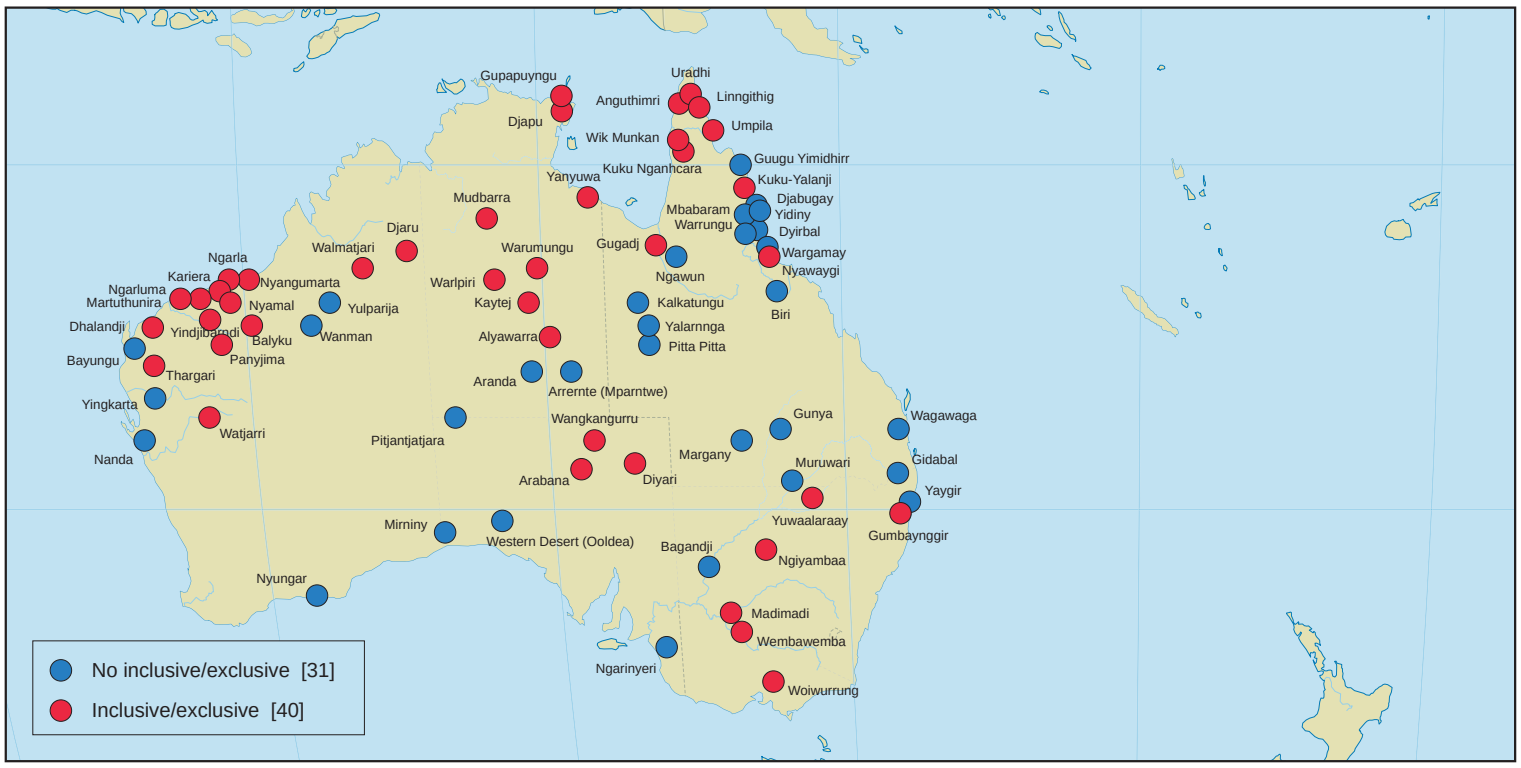
Values for Map 39A Inclusive/Exclusive Forms in Pama-Nyungan

The Pama-Nyungan stock covers Australia almost completely. Only in northern Australia are there languages which are not part of this stock, commonly called non-Pama-Nyungan (these languages are not included in this map). The Pama-Nyungan languages originally did not have an inclusive/exclusive opposition, although many Pama-Nyungan languages have developed it (Dixon 1980: 334–6).

Concerning the distribution of the inclusive/exclusive distinction among the Pama-Nyungan languages, Dixon (1980: 335) claimed that "there is no regularity to the distribution – languages of both types are found in every quarter of the continent." However, as can be seen from the distribution in Map 39A, there are clear areas with languages that have an inclusive/exclusive distinction, and areas without. First, and most prominent, the whole area that borders on the non-Pama-Nyungan languages has developed an inclusive/exclusive distinction. In Western Australia, this area extends roughly to the 22nd parallel of latitude, as claimed by O'Grady et al. (1966: 104–5). At some time in the past, this area extended through central Australia downwards into southern Australia; currently, however, this connection has been interrupted because the Central and Southern Arandic languages have recently lost the inclusive/exclusive distinction (Koch 1997). Southeastern Australia appears to be a separate area with an inclusive/exclusive distinction. However, the

situation is not completely clear, partly because there is not that much known about these languages.

The areal patterns are stronger than the genetic bonds among the Pama–Nyungan languages. For example, most languages of the Paman family in Cape York have an inclusive/exclusive distinction, except for the two southernmost languages Ngawun and Mbabaram; these border on the Maric and Galgadungic families, which do not have an inclusive/exclusive distinction.



Map 39A Inclusive/Exclusive Forms in Pama-Nyungan

40 Inclusive/Exclusive Distinction in Verbal Inflection

1 Introduction

The inclusive/exclusive opposition is a distinction which marks two different forms that are both to be translated into English as *we*. The inclusive 'we' is used if the addressee is included in the reference; it means something like 'you and I (and possibly others)'. In contrast, the exclusive 'we' is used if the addressee is *not* included in the reference. This form of 'we' means something like 'I and others (not you)'.

Map 40 shows the distribution of the inclusive/exclusive distinction in verbal inflection. This map is related to the previous Map 39, which shows the distribution of the inclusive/exclusive distinction in independent pronouns. However, there are languages that do not show the same inclusive/exclusive patterning in their independent pronouns and in their verbal inflection. About 50 per cent of the languages investigated here show the same type in both situations. Another 40 per cent simply have no verbal inflection for person at all, so the question as to any inclusive/exclusive distinction becomes irrelevant. A final 10 per cent show different marking in pronouns and verbal inflection. This is exemplified by Ngiti (Nilo-Saharan; Democratic Republic of Congo), as shown in (1). The independent pronouns distinguish *ma* ('I') from exclusive *mà* and from inclusive *alɛ̀* (tone is a distinctive feature here). However, the verbal prefixes do not distinguish between 'I' and exclusive 'we', which are both marked by a prefix *m-*.

(1) Ngiti (Kutsch Lojenga 1994: 220–1)

a. *ma* *m-òdzɨ́*
1SG.PRON 1-cry
'I cry.'

b. *mà* *m-òdzɨ́*
EXCL.PRON 1-cry
'We (exclusive) cry.'

c. *alɛ̀* *k-òdzɨ́*
INCL.PRON INCL-cry
'We (inclusive) cry.'

2 Definition of values

The definitions of the values that are distinguished in this map are almost identical to the definitions of the values used in the previous map:

1.	No person marking at all	70
2.	'We' and 'I' identical	12
3.	No inclusive/exclusive opposition	79
4.	Only inclusive differentiated	9
5.	Inclusive and exclusive differentiated	30
	total	200

The reader is referred to the text that accompanies the previous map for a detailed description of the various values. There are two small—but significant—differences in the definitions as used for this map.

First, when the inflection is distributed over two different morphological "slots" then I have only included the marking as found in the one slot that marks at least an opposition between 'I' and 'you'.

In most cases, the other slot marks plural number, which I have ignored in this map. This is exemplified in (2) by Passamaquoddy-Maliseet (Algonquian; Maine and New Brunswick). In this language, there are person/number prefixes and suffixes. Together, these affixes distinguish 'I' from exclusive and from inclusive. However, the suffixes turn out only to appear in the plural, so they look like plural markers (which I ignore). In the prefixes, there is no distinction between the marking for 'I' and the marking for exclusive, which are both marked with the prefix *n-*. There is a different prefix *k-*, which marks for inclusive. This implies that this language is classified as "only inclusive differentiated".

(2) Passamaquoddy-Maliseet (Leavitt 1996: 9)

a. *n-tíhin*
1-have
'I have it'

b. *n-tíhin-èn*
EXCL-have–1PL
'We (exclusive) have it.'

c. *k-tíhin-èn*
INCL-have–1PL
'We (inclusive) have it.'

Second, on this map, white means that there is no person marking at all in the verbal inflection of the language. For example, a language like Mandarin does not have any person marking at all on its verbs, so it is marked white here.

This definition of the value "no person marking" is closely related to value one from Chapter 102. However, the definitions are slightly different, which causes some languages to be coded differently on the two maps. In some languages, the marking of person is neither an obligatory part of the verb (as in Latin), nor is it completely absent (as in Mandarin). Such in-between cases represent stages in a transition from no person inflection towards full person inflection. For the present typology, a (rather arbitrary) division of this continuum had to be made. The basic criterion that has been used is whether the person marking, when it is marked, occurs bound onto the main verb or not. This implies that languages with so-called clitic pronouns are included if the pronouns cliticize onto the main verb. This is exemplified in (3) by Yagua (Peba-Yaguan; Peru). The independent pronoun *ráy* in (3a) is identical to the verbal proclitic in (3b). These proclitics are not obligatorily used. However, when they appear, they cliticize onto the main verb. Such verbal clitics are taken as cases of verbal inflection in this map.

(3) Yagua (Payne and Payne 1990: 367, 370)

a. *ráy* *juváay*
1SG.PRON do/work/make
'I'm working.'

b. *ray=pú́ú́chiy* *sa-dee-tu*
1SG=lead/carry 3SG-child-F
'I carry/lead his daughter.'

However, there are also languages that have clitic pronouns which do not cliticize onto the verb, but occur in a fixed position in the syntactic structure. For example, in Ngiyambaa (Pama-Nyungan; New South Wales, Australia; Donaldson 1980) pronouns can optionally cliticize onto the first word of the sentence, whatever that word may be. Such clitics are not included in this map as verbal inflection, so this language is coded as "no person marking".

A second type of person marking that is not included as inflectional marking is exemplified by Hindi in (4). In Hindi, the person features are marked on a verbal auxiliary, not on the main verb. Languages which consistently use such periphrastic constructions are coded as "no person marking" on this map.

(4) Hindi (McGregor 1995: 20)

a. *maiṁ* *calt-ā* *hūṁ*
1SG.PRON go-SG 1SG.AUX
'I go.'

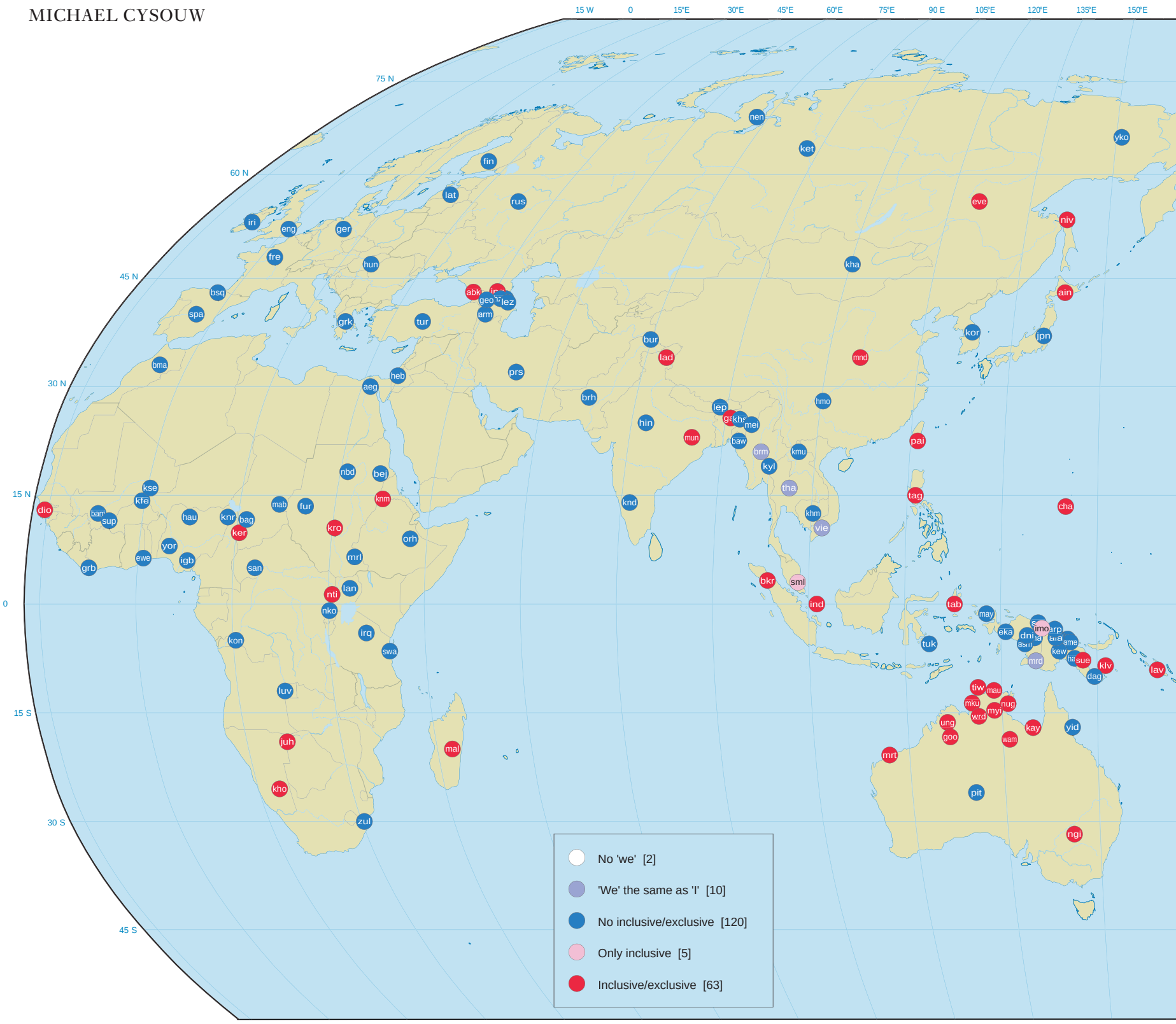
b. *ham* *calt-e* *haiṁ*
1PL.PRON go-PL 1PL.AUX
'We go.'

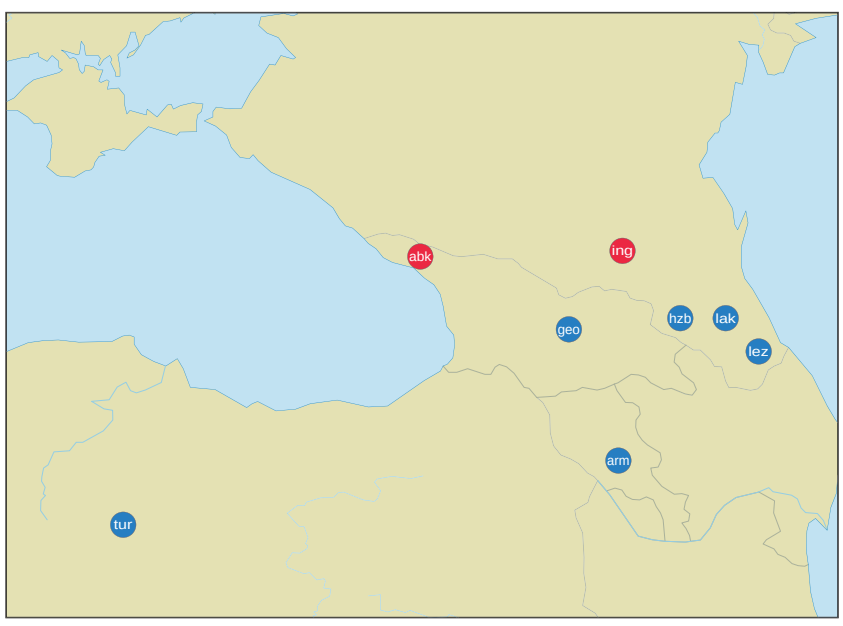
3 Geographical distribution

The inclusive/exclusive distinction in verbal inflection is uncommon in Africa and Eurasia; only sporadic cases are attested. In contrast, it is regularly attested in the Pacific. A group of Austronesian languages around New Guinea have an inclusive/exclusive distinction in their verbal inflection. However, the languages from mainland New Guinea itself do not have the distinction on verbs. The non-Pama-Nyungan languages in northern Australia constitute another clear region in which the languages have an inflectional inclusive/exclusive distinction. The Americas are particularly interesting because all five values are attested frequently, including those values that are rare worldwide. However, areally there does not appear to be any regularity within the Americas. All five values occur throughout the continent without any typologically uniform areas.

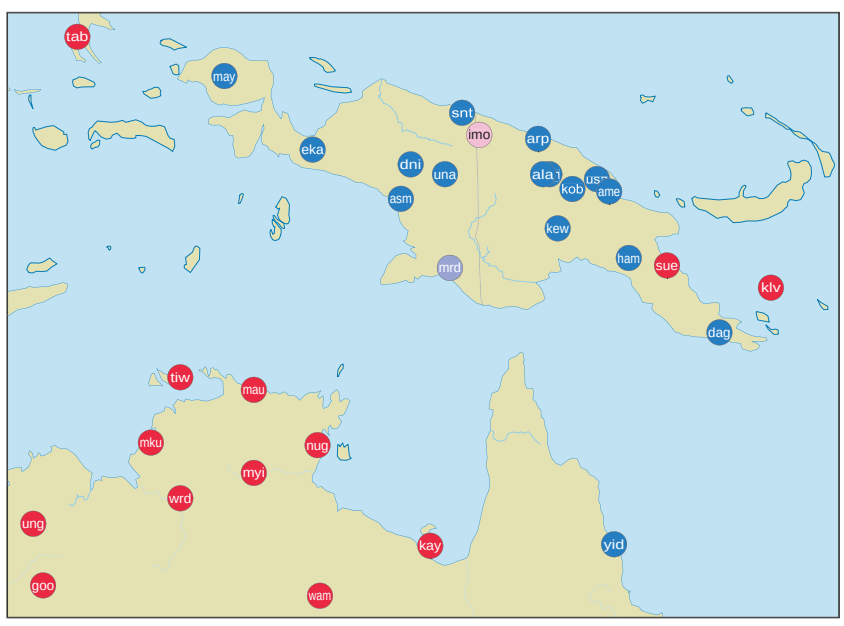
39 Inclusive/Exclusive Distinction in Independent Pronouns

MICHAEL CYSOUW

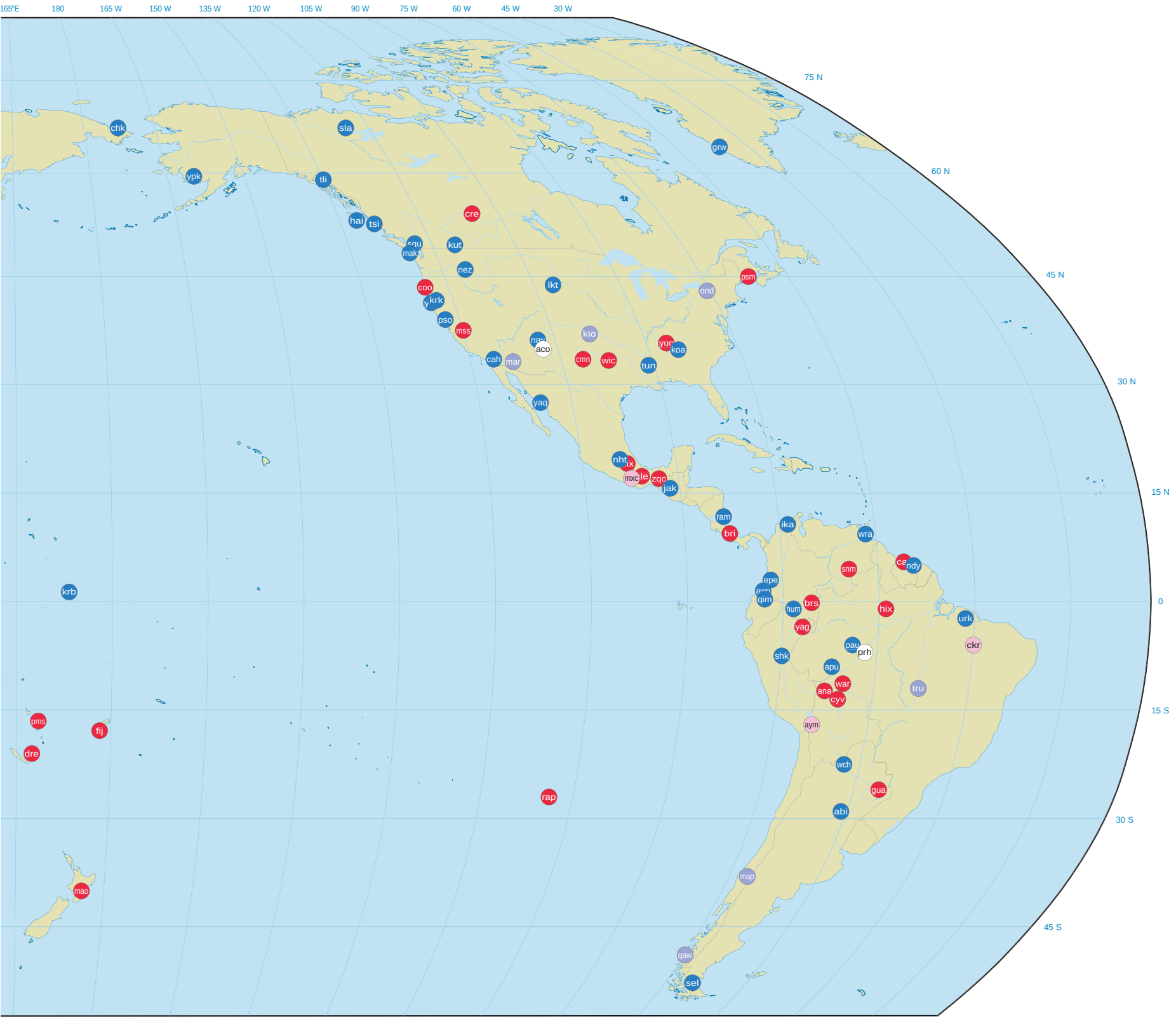




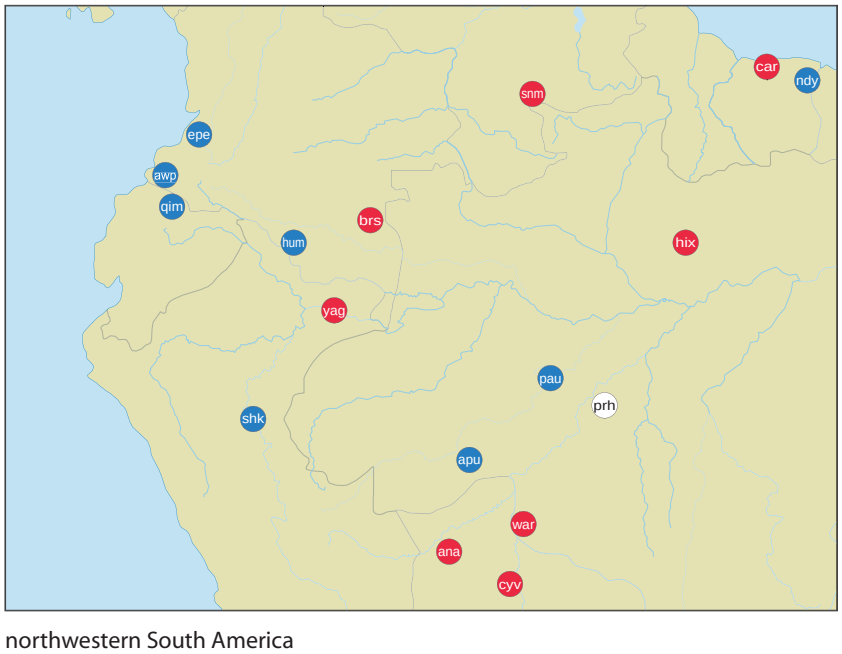
Caucasus and adjacent areas



New Guinea and northern Australia



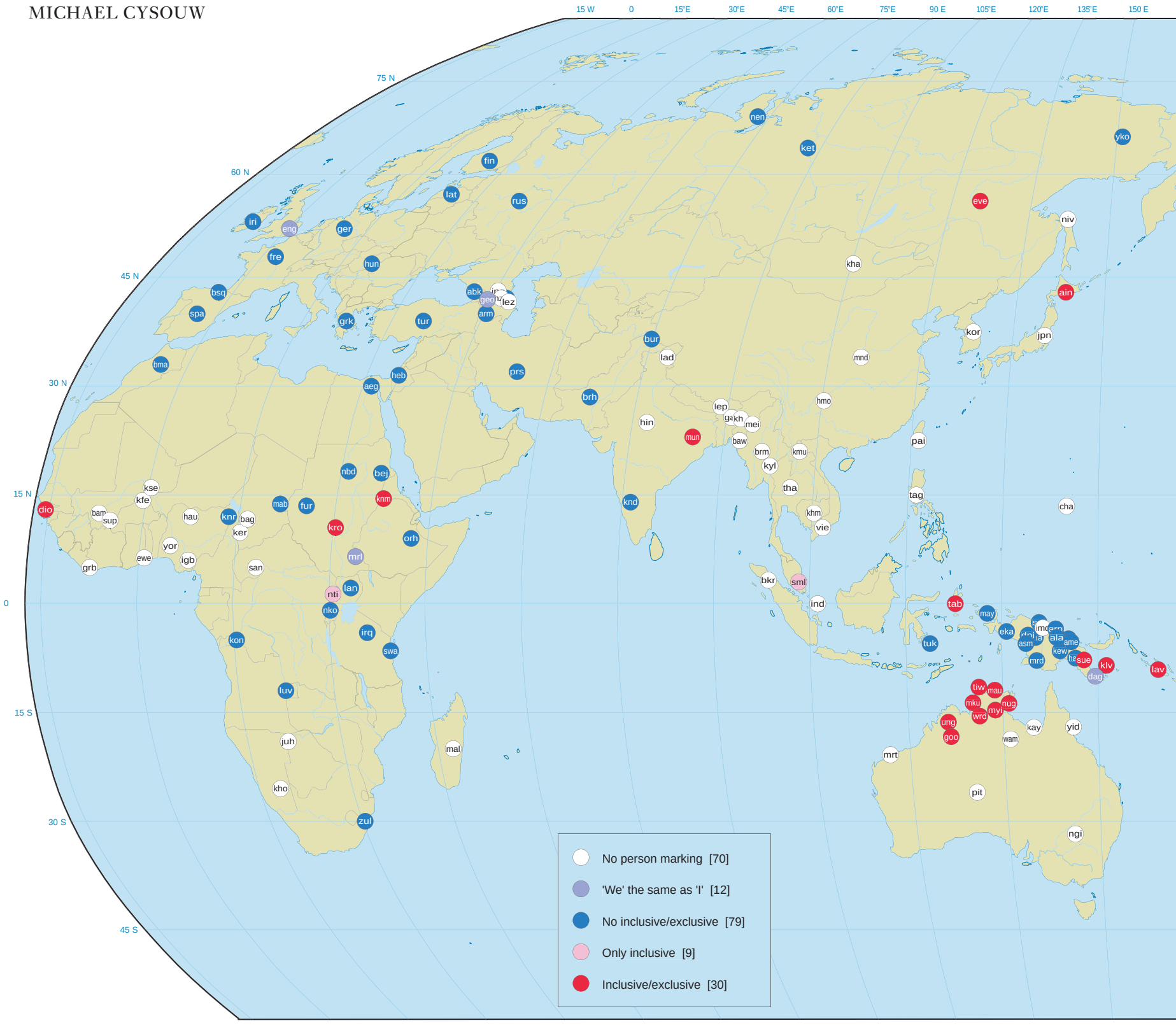
southern Mexico and Central America

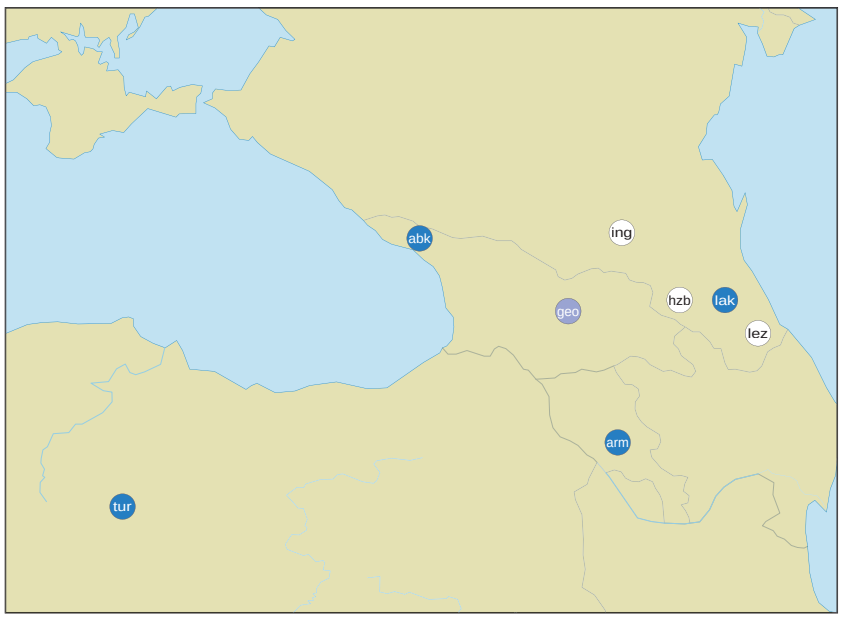


northwestern South America

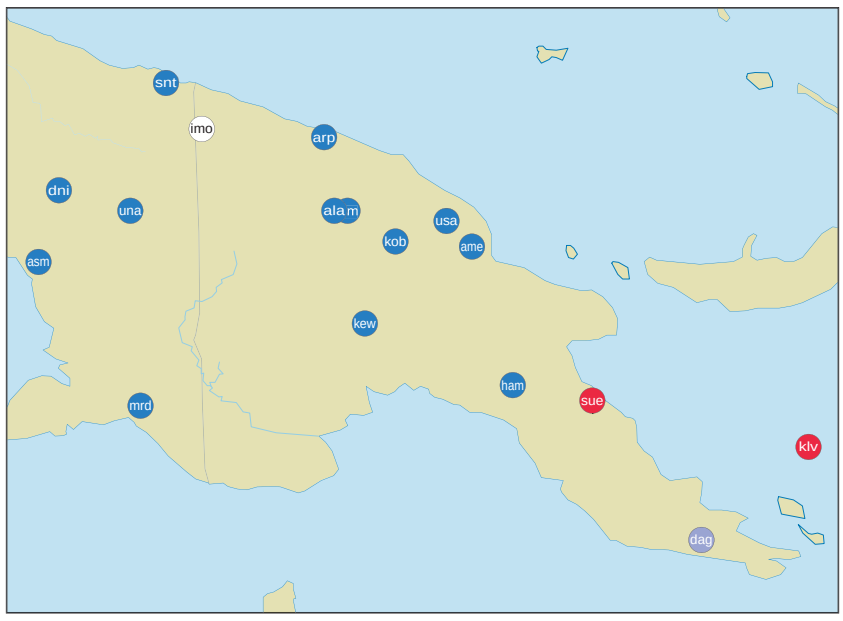
40 Inclusive/Exclusive Distinction in Verbal Inflection

MICHAEL CYSOUW

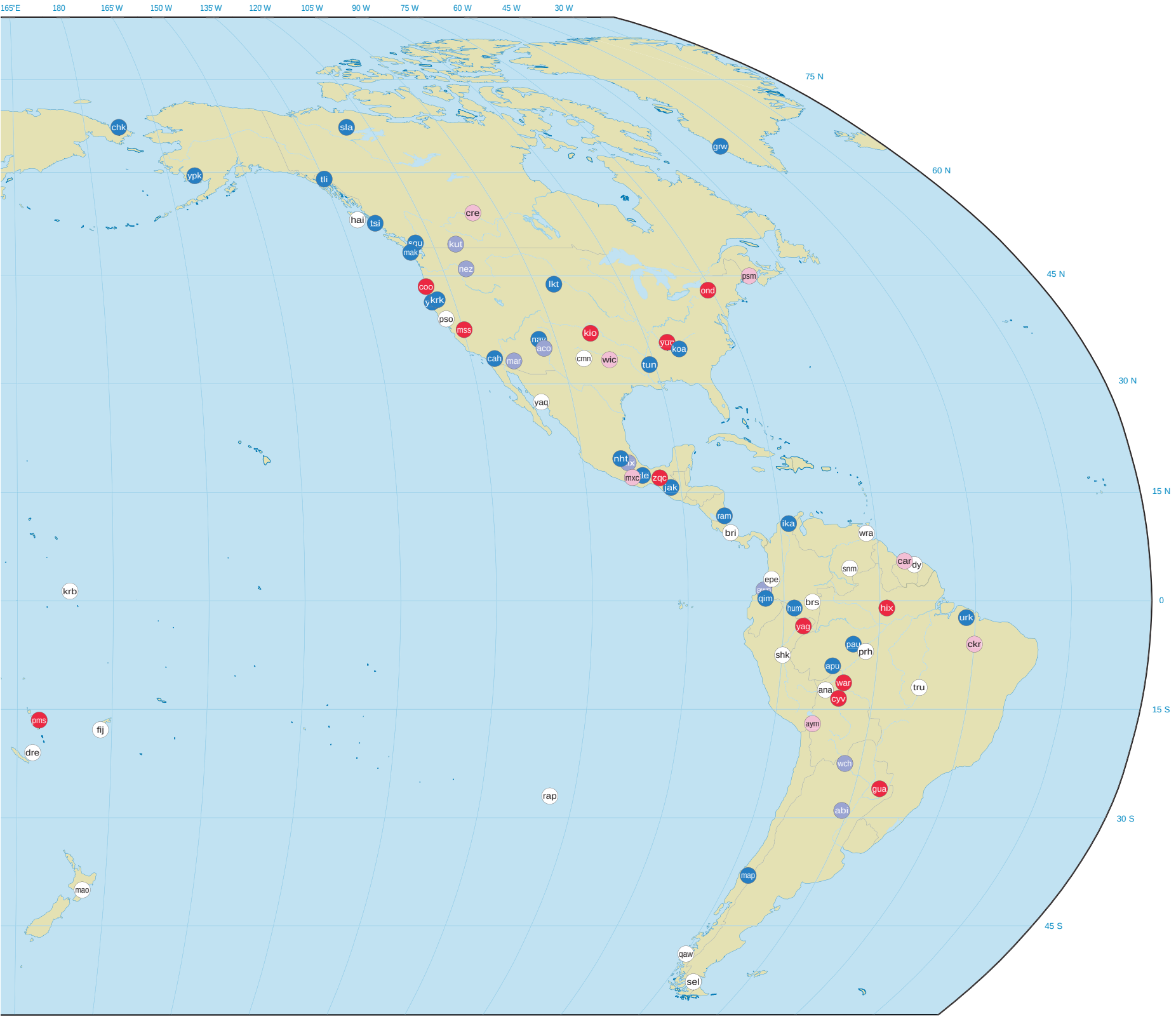


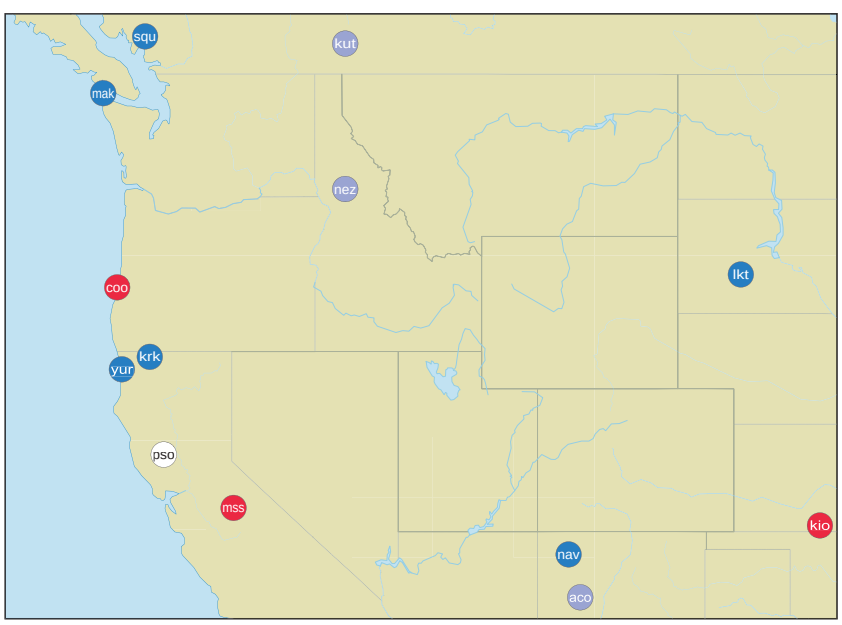


Caucasus and adjacent areas

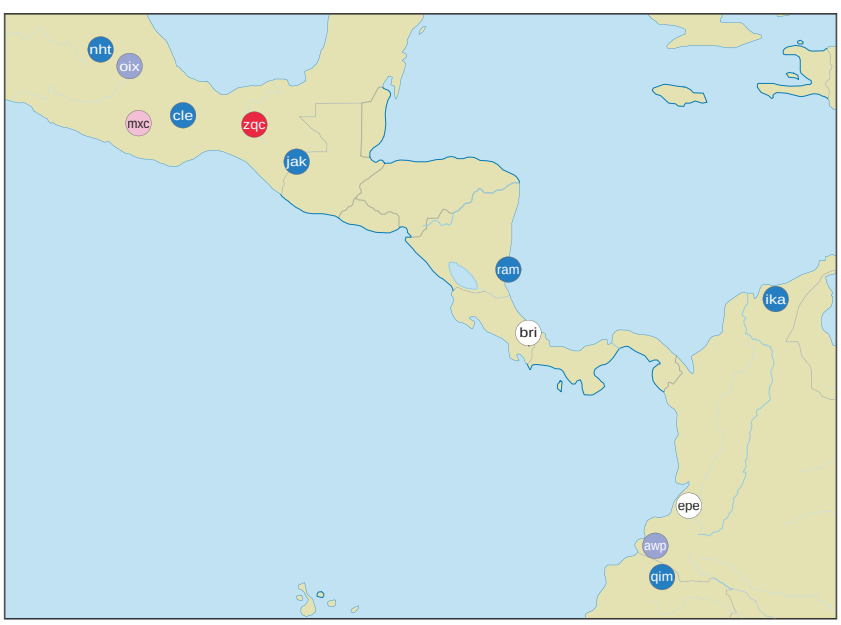


eastern New Guinea





western United States



from southern Mexico to Columbia

41 Distance Contrasts in Demonstratives

HOLGER DIESSEL

1 Defining the values

Demonstratives are deictic expressions such as English *this* and *that*. They indicate the relative distance of a referent in the speech situation vis-à-vis the **deictic centre**. The deictic centre, which is also called the **origo**, is roughly equivalent to the speaker's location at the time of the utterance. When *this* and *that* are used contrastively, *this* denotes a referent in relative proximity to the deictic centre and *that* denotes a referent at a greater distance:

(1) ***This*** *is my pen and* ***that*** *one is yours.*

The English demonstratives express a simple distance contrast. However, the demonstratives of other languages are often more specific. They may indicate not only the relative distance of the referent to the deictic centre but also whether the referent is visible or out of sight, at a higher or lower elevation, uphill or downhill, upriver or downriver, or moving toward or away from the deictic centre (see Diessel 1999: ch. 3).

This map shows the number of **distance contrasts** that are expressed in adnominal demonstratives; the other deictic features are not shown on the map. That is, if a language employs, for instance, two demonstratives for proximal and distal referents and a third demonstrative for referents out of sight, it has been classified as a two-term language, i.e. the demonstratives for referents out of sight have been disregarded.

The map is restricted to **adnominal demonstratives** because in some languages pronominal and adnominal demonstratives do not express the same distance contrasts (see Chapter 42 for a detailed discussion of these two kinds of demonstratives). For instance, in Tongan (Polynesian), adnominal demonstratives express a two-way contrast between proximal and distal referents whereas pronominal demonstratives differentiate between referents that are near the speaker, near the hearer, and away from both speaker and hearer. In addition, Tongan has a neutral demonstrative that is only used pronominally (see Table 1).

Table 1 Demonstratives in Tongan (Churchward 1953: 150–2)

Pronominal		**Adnominal**	
Neutral	*e*		
Near Speaker	*eni*	Proximal	*ni*
Near Hearer	*ena*	Distal	*na*
Away from Speaker and Hearer	*ia*		

Interestingly, if adnominal and pronominal demonstratives do not show the same distance contrasts, which is relatively uncommon, pronominal demonstratives tend to express a larger number of distance contrasts than adnominal demonstratives.

On the map, the distance contrasts of adnominal demonstratives are represented by five values:

1. No distance contrast	7
2. Two-way contrast	127
3. Three-way contrast	88
4. Four-way contrast	8
5. Five (or more)-way contrast	4
total	234

Languages that do not indicate a distance contrast are languages in which adnominal demonstratives are **distance neutral**. An example of such a language is German. Modern German has two adnominal demonstratives that precede a coreferential noun: *dieser* and stressed *der, die, das.* Note that *dieser* does not contrast with *jener* as in older stages of German; like *der, die,* and *das*, *dieser* is synchronically distance neutral.

Since the German demonstratives are deictically noncontrastive, one could argue that they are definite articles rather than demonstratives. However, for the purposes of this map German *dieser* and stressed *der, die, das* have been classified as demonstratives because of their pragmatic functions. Like distance-marked demonstratives, these expressions are commonly used to focus the hearer's attention on entities in the surrounding situation, which is not what speakers usually do with definite markers.

German may express a deictic contrast between two referents by adding the **adverbial demonstratives** *hier* 'proximal' and *da* 'distal' to the neutral forms:

(2) German
Das *Bild* ***hier*** *gefällt mir besser als* ***das*** ***da***.
DEM picture here like me better than DEM there
'I like this picture better than that one (over there).'

The same strategy is used in other languages in which adnominal demonstratives are distance neutral (e.g. French *ce N-ci* versus *ce N-là*). In general, it seems that all languages have at least two adverbial demonstratives (or deictic particles) that indicate a distance contrast between two referents or locations, and if the adnominal demonstratives are distance neutral, the adverbial demonstratives can be added to the neutral forms when it is necessary to differentiate between two referents.

In languages in which adnominal demonstratives express a two-way distance contrast, the demonstratives refer to proximal and distal referents. Note, however, that in most of these languages demonstratives can also be used noncontrastively. For instance, in Modern Hebrew *ze* indicates a proximal referent when it is used in contrast to *hu* 'distal'; but *ze* is also used in noncontrastive situations, where it is distance-neutral.

Other two term languages employ a special neutral term in addition to distance-marked demonstratives. For instance, in Lithuanian *šìs* 'proximal' and *anàs* 'distal' are deictically contrastive and *tàs* is distance neutral. Languages like Lithuanian have been classified as two-term languages because the neutral demonstrative does not add another distance contrast to the system.

2 Distance- and person-oriented systems

Languages in which adnominal demonstratives express a three-way contrast can be divided into two types: (1) **distance-oriented systems**, in which all demonstratives indicate the relative distance between the referent and the deictic centre, and (2) **person-oriented systems**, in which one of the demonstratives denotes a referent in the proximity of the hearer (see Anderson and Keenan 1985). Hunzib (Daghestanian; eastern Caucasus) has a distance-oriented system including three adnominal demonstratives that indicate three different locations on a distance scale (see Table 2), and Japanese has a person-oriented system in which one of the three demonstratives refers to an entity near the hearer (see Table 3).

Table 2 Demonstratives in Hunzib (van den Berg 1995: 61)

	adnominal demonstratives
Proximal	*bəd*
Medial	*bəl*
Distal	*əg*

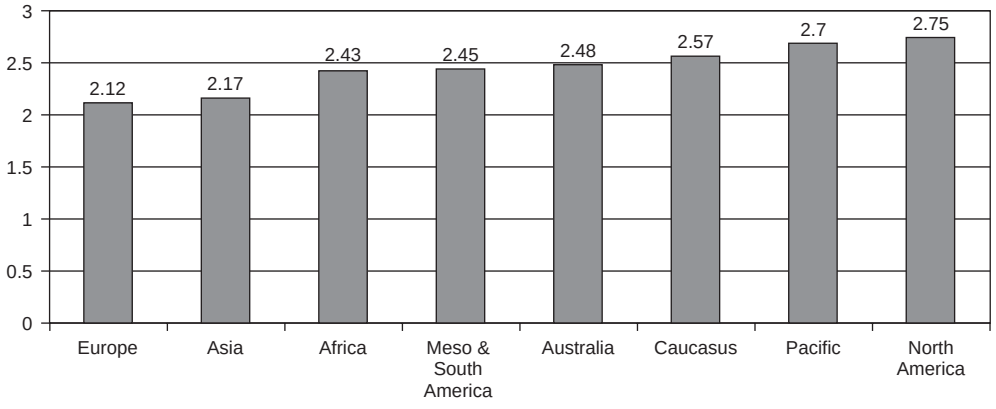


Figure 1 Mean number of distance contrasts in demonstratives

Table 3 Demonstratives in Japanese (Kuno 1973: 27)

	adnominal demonstratives
Near speaker	*kono*
Near hearer	*sono*
Away from speaker and hearer	*ano*

Languages having more than three adnominal demonstratives are almost always person-oriented; that is, they usually employ a special term for referents near the hearer. An example of a four-term system comes from Hausa (see Table 4).

Table 4 Demonstratives in Hausa (Wolff 1993: 119–20)

	adnominal demonstratives
Near speaker	*nân*
Near hearer	*nan*
Away from speaker and hearer	*cân*
Far away from speaker and hearer	*can*

3 Geographical distribution

The vast majority of the world's languages employ two or three distance-marked demonstratives: 54.3 per cent of all languages shown on the map have adnominal demonstratives that express a two-way contrast between proximal and distal referents and 37.6 per cent have adnominal demonstratives that differentiate between three referents. The majority of the latter have a distance-oriented system; that is, they distinguish between three different locations on a distance scale. Person-oriented systems seem to be less common: about two thirds of the three-term languages shown on the map have been classified as distance-oriented systems in reference grammars, and only one third as person oriented.

One-term systems, in which adnominal demonstratives are distance neutral, and larger systems having four, five, or more adnominal demonstratives, are cross-linguistically uncommon. Together they account for only 8 per cent of the languages shown on the map. In contrast to the three-term systems, the four- and five-term systems are almost always person oriented.

If we look at the areal distribution of the various distance systems, we find that two- and three-term systems are prevalent throughout the world. However, the occurrence of smaller and larger systems is restricted to certain areas. One-term systems occur quite frequently in Africa (Supyire, Koyraboro Senni, Koromfe, Kera), but they are also found in Europe (German, French) and Meso-America (Mam). Four- and five-term systems occur in three regions: North America (Koasati, Maricopa, Navajo, Tlingit, Quileute), Africa (Hausa, Iraqw, Somali, Malagasy), and the Pacific region (Kambera, Paakantyi, Asmat). Figure 1 shows the mean number of distance contrasts that are expressed by adnominal demonstratives in various parts of the world.

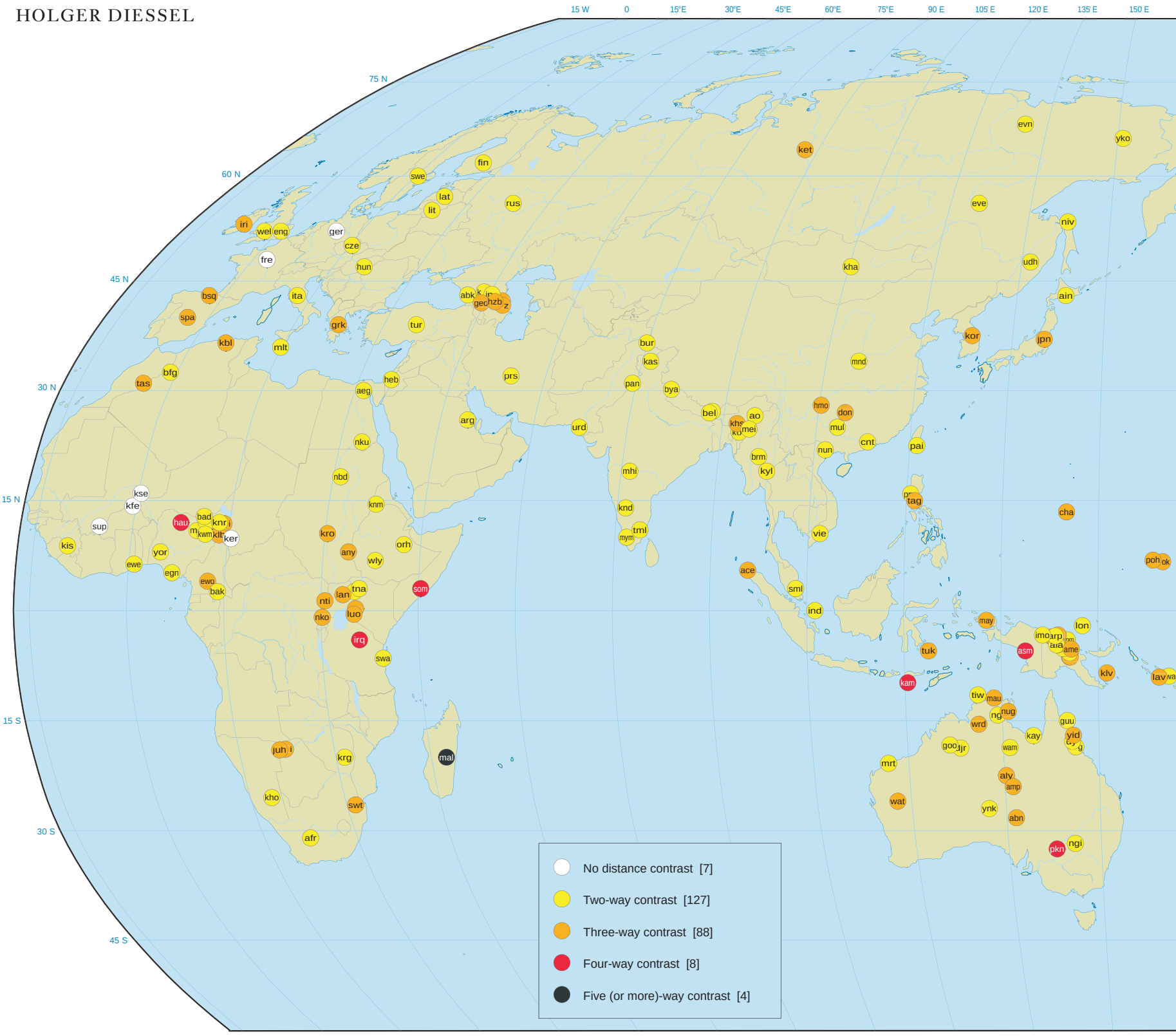
As can be seen in this figure, there are two areas where adnominal demonstratives express relatively few distance contrasts, Europe and Asia. Especially high is the number of distance contrasts in North America and in the Pacific. However, overall the variation is relatively small across the different regions. It ranges from a mean distance contrast of 2.12 in Europe to 2.75 in North America; that is, the difference between the area with the fewest distance contrasts and the area with the most distance contrasts amounts to about 30 per cent.

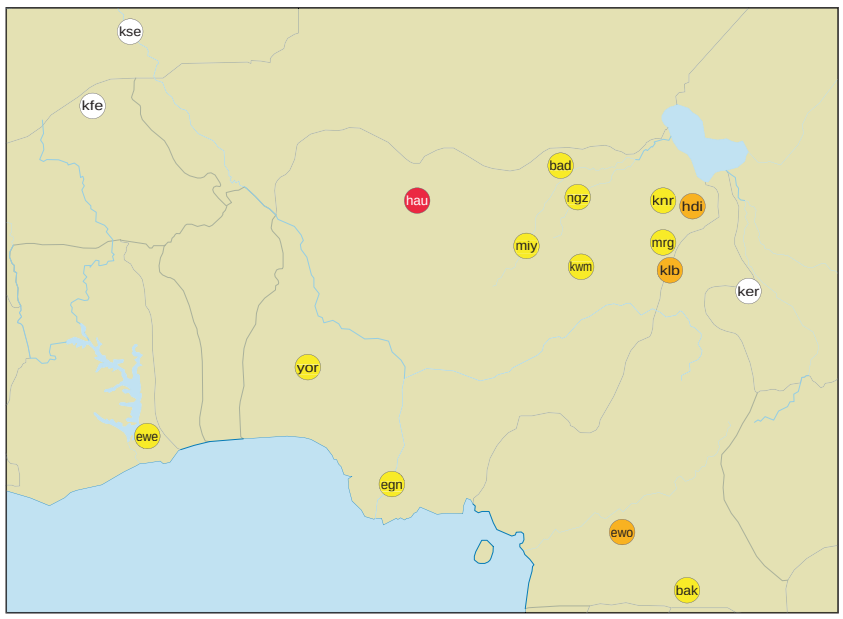
4 Theoretical issues

The interpretation of the distance contrasts that are encoded by demonstratives crucially depends on the conceptualization of the deictic centre. In a distance-oriented system, the deictic centre is basically the same for all demonstratives. However, in a person-oriented system one has to distinguish between two different origos (Diessel 1999: ch. 3). The origo of the first and second terms is defined by the location of the speaker, excluding the location of the hearer, whereas all other terms relate the referent to a domain including both speaker and hearer. In other words, in a person-oriented system the deictic centre is conceptualized in two different ways: it is the domain of the speaker alone (excluding the location of the hearer) for the first and second terms, and it is the common domain of speaker and hearer for the distal term(s). Thus, from a conceptual perspective, a person-oriented system with three distance terms is a variant of a distance-oriented system with two distance terms, in that the demonstratives of such a three-term system never distinguish between more than two referents, namely the deictic centre (which is either the location of the speaker or the common domain of speaker and hearer) and the referent. Similarly, a person-oriented system with four demonstratives can be seen as a conceptual variant of a distance-oriented system with three distance terms. What this means is that with very few exceptions (notably the languages that have more than four adnominal demonstratives), languages do not invoke more than three different locations on the distance scale. For the great majority of the world's languages, three appears to be the upper limit; there is very little cross-linguistic variation in this domain.

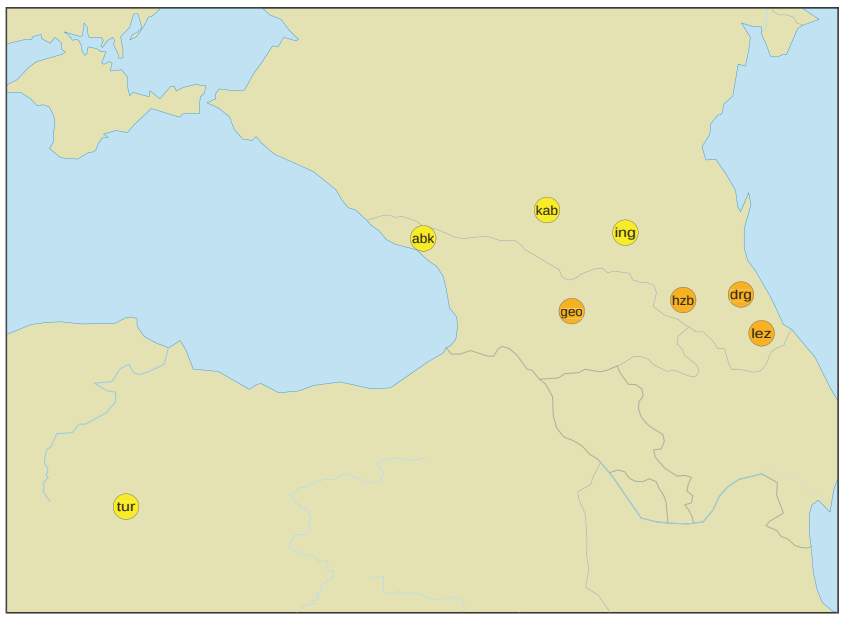
41 Distance Contrasts in Demonstratives

HOLGER DIESSEL

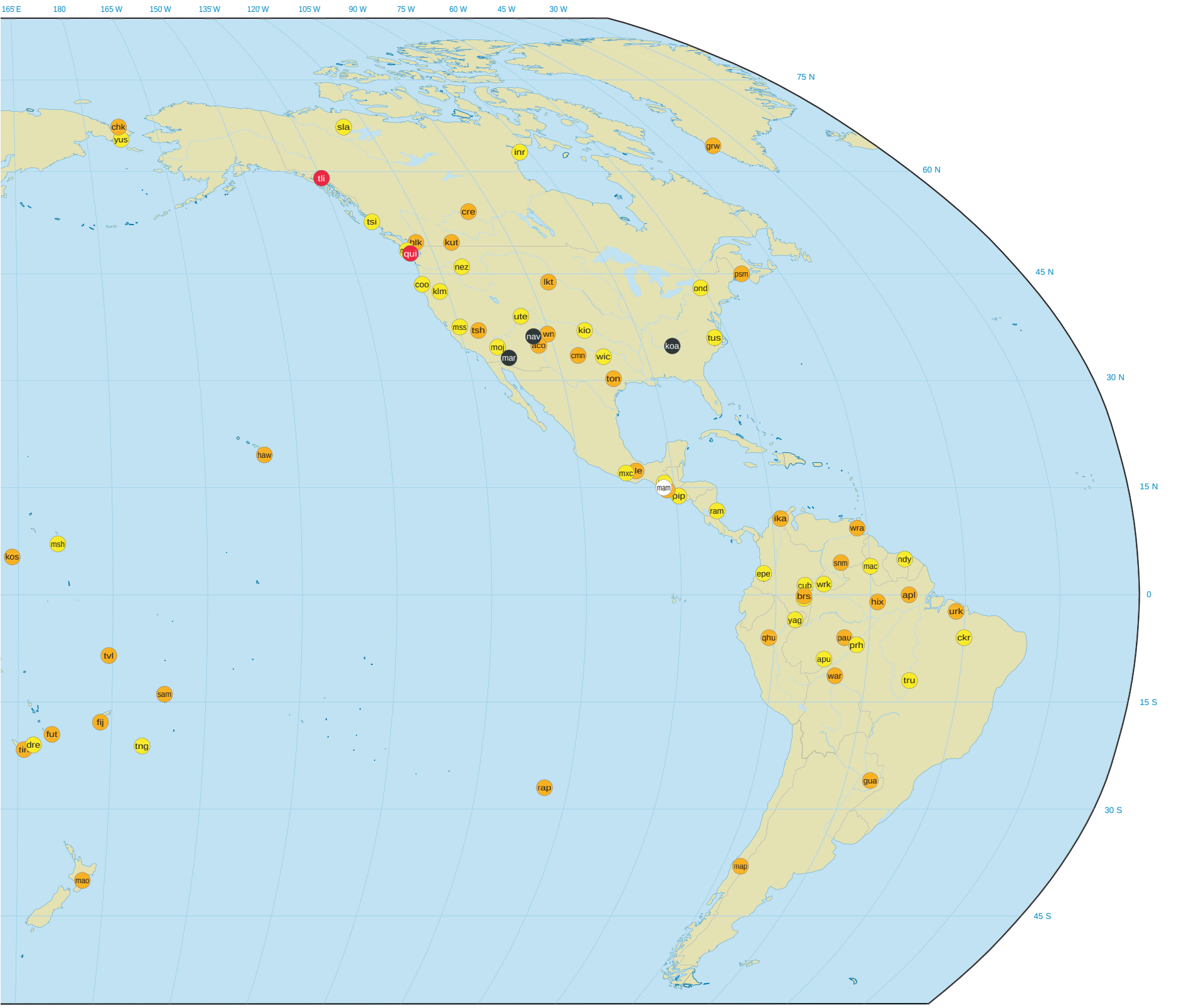


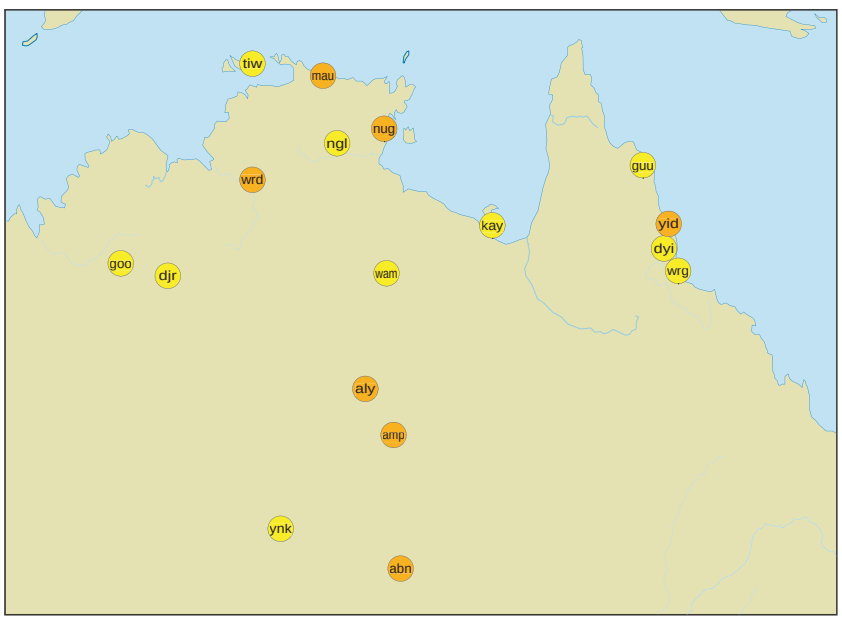


Nigeria and surrounding areas

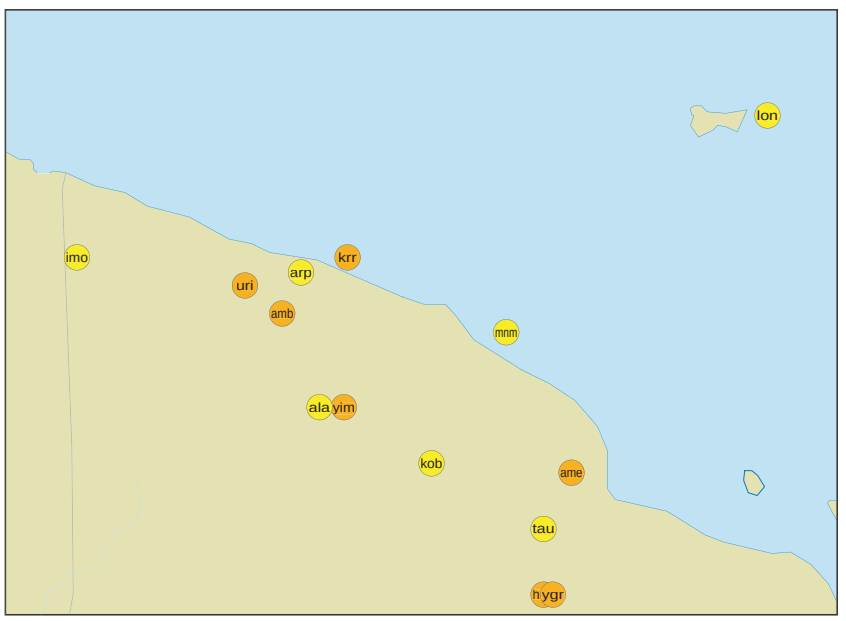


Caucasus and adjacent areas





northern Australia



eastern New Guinea

42 Pronominal and Adnominal Demonstratives

HOLGER DIESSEL

1 Defining the values

Demonstratives are commonly divided into **pronominal demonstratives**, which substitute for a noun (phrase), and **adnominal demonstratives**, which accompany a coreferential noun. In English, pronominal and adnominal demonstratives have the **same forms**: *this* and *that* may function as independent pronouns, as in (1a), or they may co-occur with a noun, as in (1b).

(1) a. *I don't like* ***that****.*
b. ***this*** *book*

In other languages, pronominal and adnominal demonstratives are often formally distinguished: they may have **different stems** or they may have **different inflectional features**. For instance, in French pronominal and adnominal demonstratives have different stems: *celui* and *celle* are used pronominally and *ce* and *cette* accompany a noun.

(2) French
Donne-moi *ce* *livre-là* *et* *garde* ***celui-ci***
give-me this book-there and keep this.one-here
pour *toi.*
for you
'Give me that book and keep this one for you.'

In Turkish, pronominal and adnominal demonstratives have the same stems; both involve the proximal demonstrative *bu* and the distal demonstrative *o*. However, pronominal and adnominal demonstratives differ in their inflectional behaviour: pronominal demonstratives are inflected for case (and number) (see (3a)), whereas the adnominal demonstratives are uninflected particles that precede an inflected noun (see (3b)).

(3) Turkish (Kornfilt 1997: 312, 315)
a. *Ali* ***bun-u*** *unut-amı-yor.*
Ali this-ACC forget-cannot-PROG
'Ali is unable to forget this.'
b. ***bu*** *gazete-yi*
this newspaper-ACC
'this newspaper'

Map 42 shows the geographical distribution of three types of languages: languages like English, in which the demonstratives have the same forms (value one); languages like French, in which the demonstratives have different stems (value two); and languages like Turkish, in which the demonstratives have different inflectional features (value three).

1. Same forms	143
2. Different stems	37
3. Different inflectional features	21
total	201

Note that demonstratives of the third type may or may not have different stems; what is crucial is that they have different inflectional features. For instance, in Tamil the pronominal demonstratives have the stems *itu* 'proximal' and *atu* 'distal', which are combined with gender-number affixes. The adnominal demonstratives, on the other hand, are expressed by *intu* 'proximal' and *antu* 'distal', which do not occur with gender-number markers. Since pronominal and adnominal demonstratives have different inflectional properties in Tamil, Tamil is classified as a language of the third type regardless of the fact that the demonstratives also have different stems.

In the great majority of languages, pronominal and adnominal demonstratives are independent words. However, adnominal demonstratives can also be clitics. Demonstrative clitics are bound forms that attach to a noun or another word in the noun phrase. Two examples from Lango (Nilotic; Uganda) are given in (4a–b).

(4) Lango (Noonan 1992: 155)
a. *gwók=**kì***
dog=this
'this dog'
b. *gwôkk* *à* *dwóŋ=**ŋì***
dog ATTR big=this
'this big dog'

In contrast to adnominal demonstratives, pronominal demonstratives are virtually always free forms (Diessel 1999: ch. 2). Languages in which adnominal demonstratives are clitics and pronominal demonstratives are independent forms are classified as languages in which pronominal and adnominal demonstratives have different stems (i.e. they have been assigned to the second type).

Some languages do not have a separate class of demonstrative pronouns. For instance, in Korean demonstratives are uninflected particles that must be accompanied by a noun. The semantic equivalent of a pronominal demonstrative in English is a noun phrase consisting of the demonstrative particles *i* 'near speaker', *ku* 'near hearer', or *ce* 'away from speaker and hearer' and a "defective noun" (Sohn 1994: 295) that indicates the type of referent (e.g. *il* 'thing/fact'):

(5) Korean (Sohn 1994: 295)
[ce *il-ul]* *nwu-ka* *mak-keyss-ni*
that thing-ACC who-NOM block-FUT-Q
'Who would be able to block that?'

There are other languages in which similar expressions are formed by combining an adnominal demonstrative with a third-person pronoun, a classifier, or some other element that functions as the head of the construction (Diessel 1999: ch. 4). Although, strictly speaking, these languages do not have a separate class of demonstrative pronouns, one can think of the construction in (5) and parallel constructions in other languages as complex pronominal forms. On the map, languages like Korean have been classified as languages in which pronominal and adnominal demonstratives have different stems (i.e. they also have been assigned to the second type).

Apart from pronominal and adnominal demonstratives, many languages employ a separate class of **identificational demonstratives**, which in many descriptions are not properly distinguished from pronominal demonstratives (Diessel 1999: ch. 4). Identificational demonstratives occur in copular and nonverbal clauses. For instance, in Pohnpeian (Oceanic; Micronesia) *me(t)* 'near speaker', *men* 'near hearer', and *mwo* 'away from speaker and hearer' are used as pronominal demonstratives in verbal clauses, whereas *ie(t)* 'near speaker', *ien* 'near hearer', and *io* 'away from speaker and hearer' function as identificational demonstratives in nonverbal clauses.

(6) Pohnpeian (Rehg 1981: 143, 150)
a. ***met*** *pahn* *mengila*
this will wither
'This will wither.'
b. ***iet*** *noumw* *naipen*
this/here your knife
'Here is your knife.'

Like Pohnpeian, Tümpisa Shoshone (Uto-Aztecan; California and Nevada) distinguishes pronominal demonstratives in verbal clauses from identificational demonstratives in nonverbal clauses. However, in Tümpisa Shoshone pronominal and identificational demonstratives have the same deictic roots, but differ in their

inflectional behaviour. The pronominal demonstratives are inflected for number and case and may take a "proximate marker" (Dayley 1989b: 136) that is prefixed to the deictic root. By contrast, the identificational demonstratives are unmarked for number, take the suffix *-sü(n)* in place of a regular case ending, and never occur in the proximate form.

(7) Tümpisa Shoshone (Dayley 1989b: 141, 145)

a. *u punikka* ***s-e-tü***
it see PROX-this-NOM
'This one saw it.'

b. ***e-sü(n)*** *nahim pungku*
this.is our.DU pet
'This is our pet.'

While identificational demonstratives can be seen as pronouns of some type, they should be distinguished from pronominal demonstratives in verbal clauses. For the purposes of this map the category of pronominal demonstratives has been restricted to demonstratives that function as arguments of verbs in verbal clauses—i.e. identificational demonstratives have been disregarded.

2 Geographical distribution

In most languages, pronominal and adnominal demonstratives are morphologically identical. In 71.1 per cent of the languages shown on the map pronominal and adnominal demonstratives have the same forms, in 18.4 per cent pronominal and adnominal demonstratives have different stems, and in 10.4 per cent they differ with regard to their inflectional features.

If we look at the geographical distribution of languages in which pronominal and adnominal demonstratives are formally distinguished, we find a concentration in two large areas. Languages in which pronominal and adnominal demonstratives differ in terms of their inflection are especially frequent in Asia and southeastern Europe. They occur in a region ranging from Turkey and the Caucasus in the west to India and Japan in the east. In the majority of these languages, adnominal demonstratives are uninflected particles that precede an inflected noun, whereas pronominal demonstratives are inflected for number and case, and less frequently also for gender. Outside of this area, there are only a few other languages on the map in which pronominal and adnominal demonstratives have different inflectional features (Somali, Tauya, Epena Pedee, and Lealao Chinantec).

Languages employing different stems are especially frequent in northern and central Africa. The majority of the Afro-Asiatic and Nilo-Saharan languages spoken in this region have demonstratives with different stems in pronominal and adnominal positions. Apart from the Afro-Asiatic and Nilo-Saharan languages, there are various other languages throughout the world in which pronominal and adnominal demonstratives have different stems. Such languages occur for instance in Europe (e.g. French), Asia (e.g. Mulao), South America (e.g. Wari'), and the Pacific region (e.g. Ambulas).

There are two large areas in which all demonstratives have the same morphological forms. One is Australia, where not a single language shown on the map differentiates between pronominal and adnominal demonstratives. The other area is North America with the exception of a small region in the Pacific North-West. The languages that are spoken in the latter region belong to the Salishan and Wakashan language families, which tend to have complex systems of articles and determiners that are formally distinct from the corresponding pronouns. In all other areas of North America, adnominal and pronominal demonstratives are expressed by the same forms.

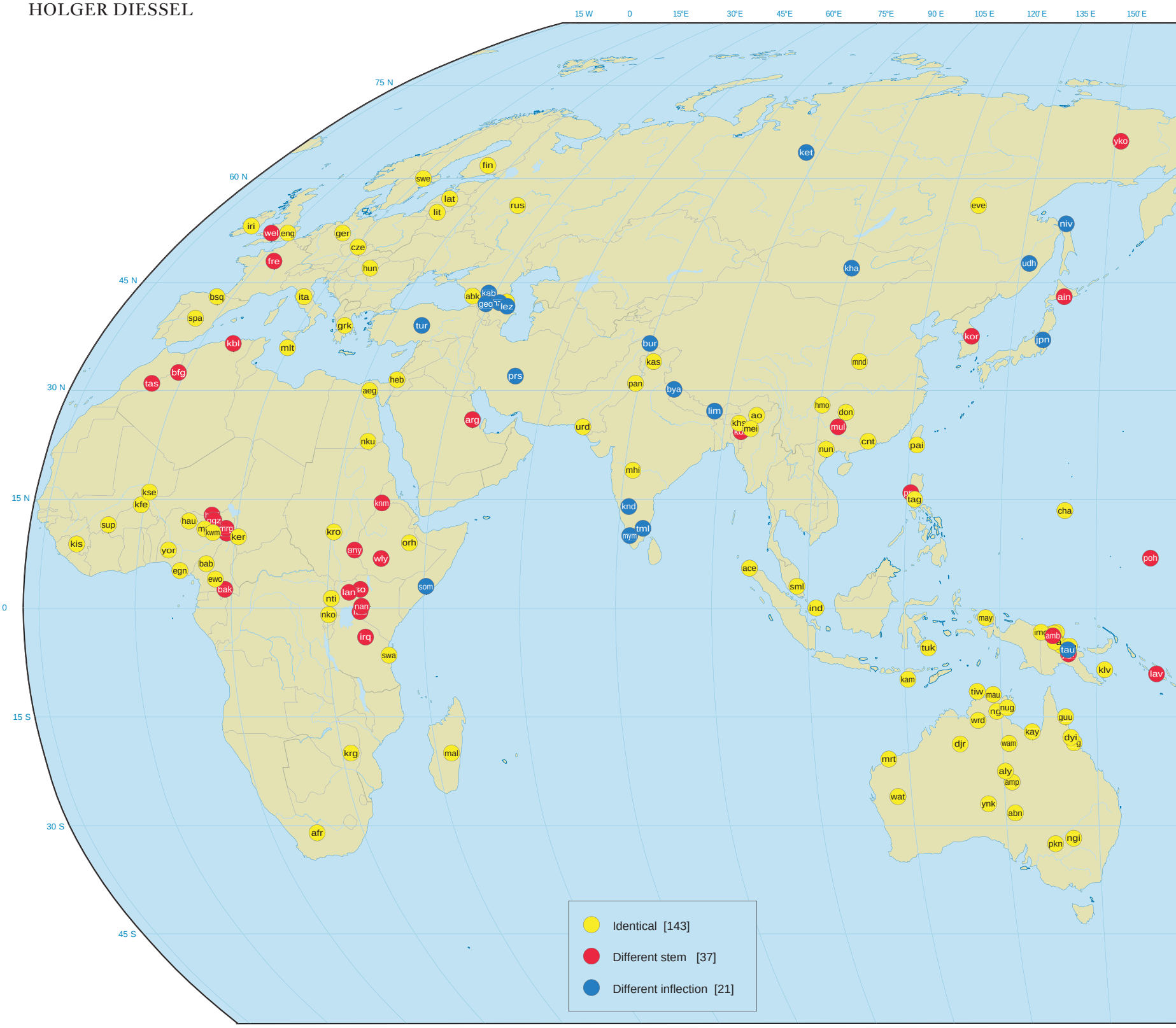
3 Theoretical issues

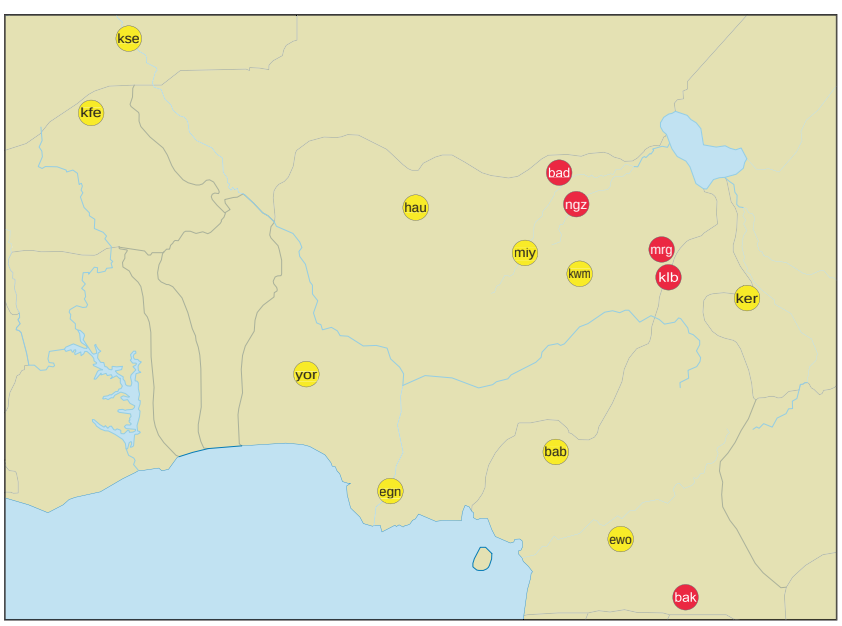
If a language employs different demonstratives in pronominal and adnominal positions, it is reasonable to assume that the demonstratives in these positions are categorially distinguished: pronominal demonstratives are independent pronouns and adnominal demonstratives are articles or determiners. However, what is the categorial status of adnominal demonstratives in languages in which pronominal and adnominal demonstratives have the same forms? Are they demonstrative pronouns, or do they function as determiners that happen to have the same forms as pronominal demonstratives?

There is good evidence that the adnominal demonstratives of many languages are independent pronouns that are only loosely adjoined to the noun in some kind of appositional structure. In particular, in languages in which the ordering of demonstrative and noun is flexible and/or discontinuous, it is reasonable to assume that adnominal demonstratives are independent pronouns in apposition to a noun (Diessel 1999: ch. 4). Languages of this sort usually belong to the nonconfigurational language type, in which words are organized in appositional constructions rather than in hierarchically structured phrases (Hale 1983). Interestingly, languages of the two areas where pronominal and adnominal demonstratives always have the same forms, Australia and North America, are well known for their nonconfigurational properties. An interesting hypothesis is thus that the pronominal and adnominal demonstratives in the Australian and North American languages are formally identical because the adnominal demonstratives of these languages are demonstrative pronouns in apposition to a noun.

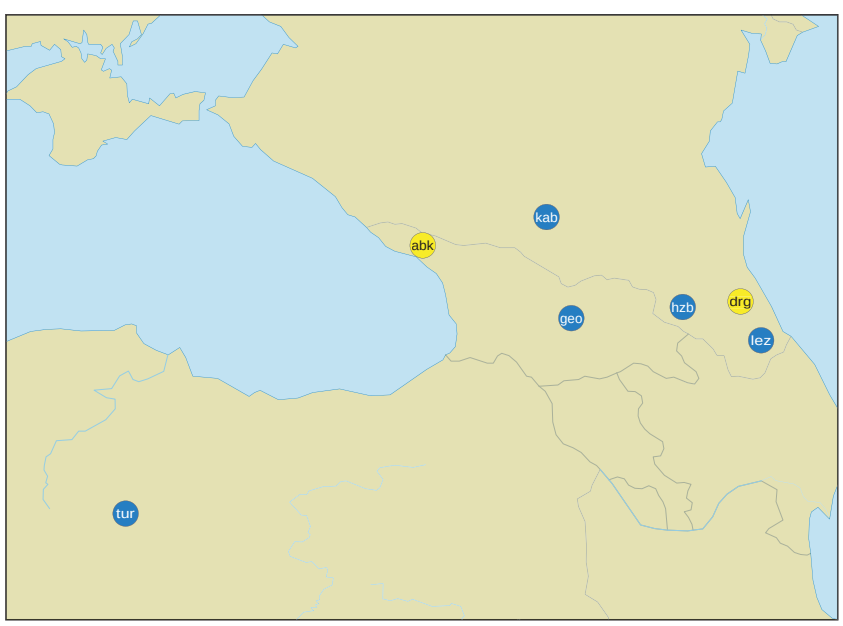
42 Pronominal and Adnominal Demonstratives

HOLGER DIESSEL

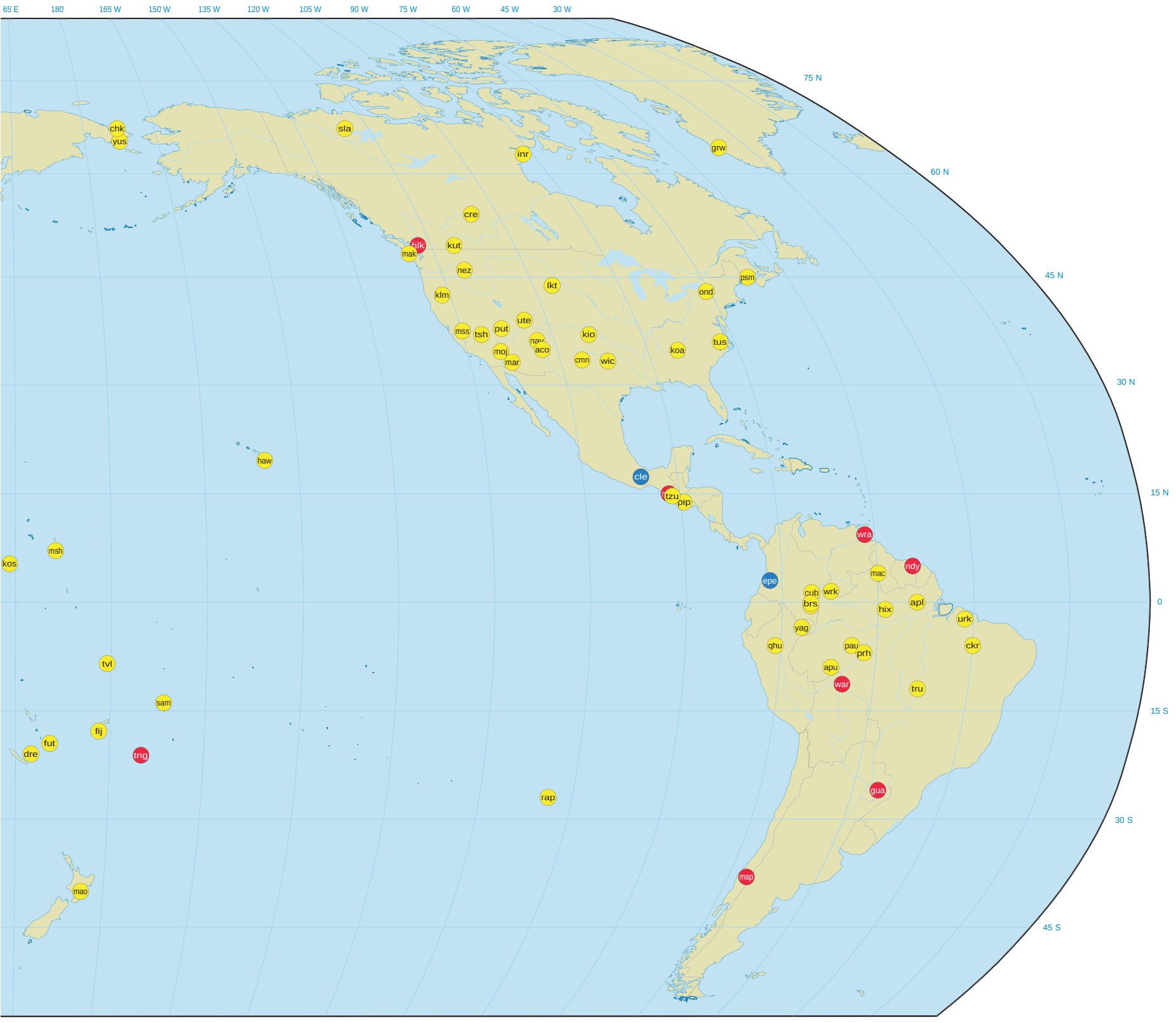


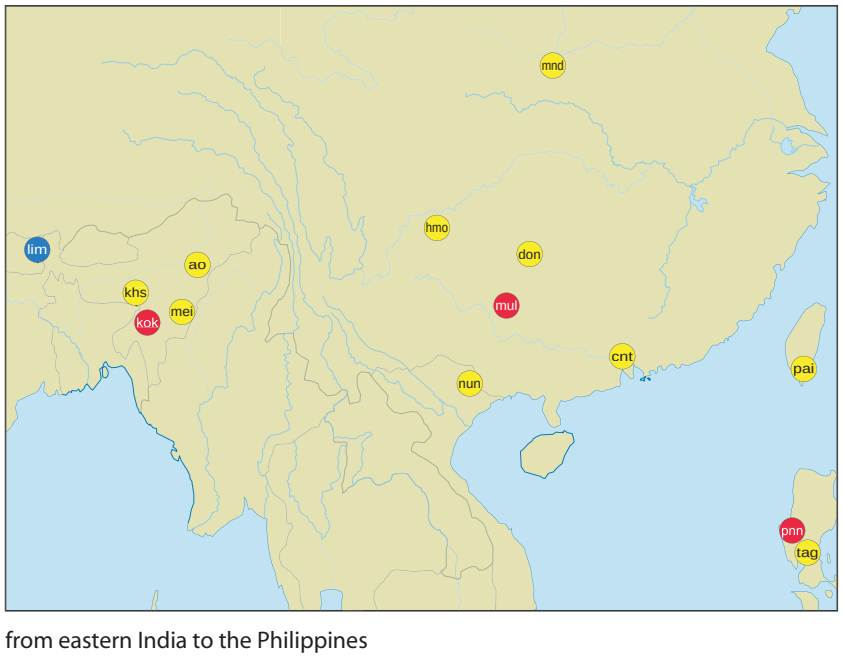


Nigeria and surrounding areas

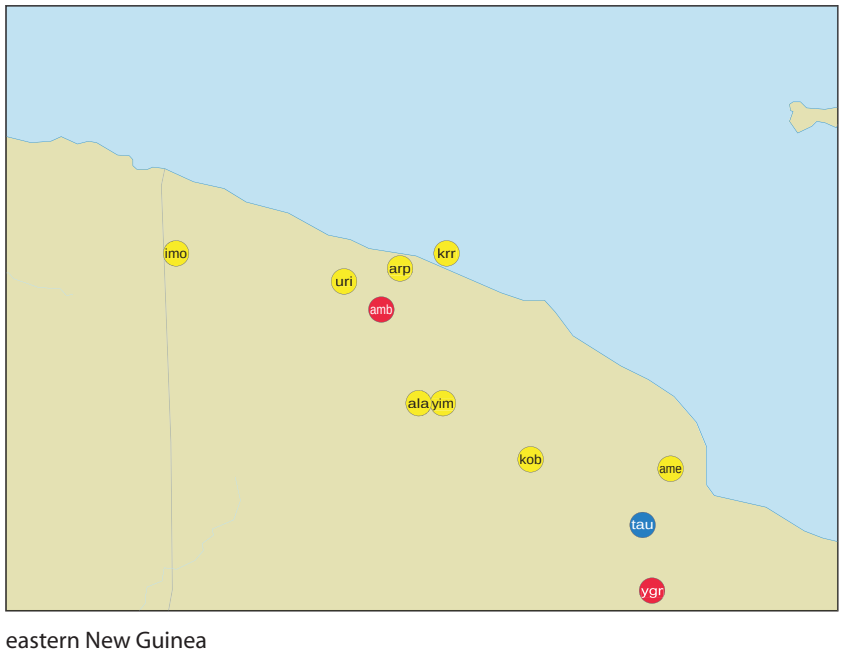


Caucasus and adjacent areas





from eastern India to the Philippines



eastern New Guinea

43 Third-Person Pronouns and Demonstratives

D. N. S. BHAT

1 Defining the values

Map 43 examines the relationship between **third-person pronouns** and **demonstratives** on the basis of data gathered from 225 languages. The primary division here is between languages in which third-person pronouns and demonstratives are unrelated and languages in which they are either identical or derivationally related. The latter languages are further divided into five different groups depending upon the type of relation that exists between the two.

Third-person pronouns and demonstratives are:

○	1. unrelated	100
●	2. related for all demonstratives	52
◆	3. related for remote demonstratives	18
■	4. related for non-remote demonstratives	14
●	5. related by gender markers	24
●	6. related for nonhuman reference	17
	total	225

1.1 No relation. Among the 225 languages examined, 100 belong to the first group, in which third-person pronouns are unrelated to demonstratives. In Ainu (Hokkaido, Japan), for example, the third-person pronoun is *sinuma*, which is unrelated to the demonstratives *tan* 'this' and *toan* 'that' (Tamura 2000: 48, 93). Similarly in Epena Pedee (Choco; Colombia), *iru* 'third person' is unrelated to the demonstratives *na* 'this' and *hã* 'that' (Harms 1994: 45, 57).

The remaining 125 languages show either identity or derivational relationship between third-person pronouns and demonstrative pronouns. These languages are divided into five different types (values 2–6), depending upon the type of affinity that obtains between the two.

1.2 Third-person pronouns related to all demonstratives. In the case of 52 of these 125 languages, the affinity derives either from the fact that any of the demonstratives can be used as third-person pronouns, or from the fact that third-person pronouns and demonstratives share the same stem. In Basque, for example, *hau* 'proximal', *hori* 'medial', and *hura* 'distal' are third-person pronouns that also function as demonstrative pronouns (Saltarelli et al. 1988: 213). Thirty-three languages (out of the 52) have demonstratives any one of which can be used as a third-person pronoun.

In the case of four languages, there is a distinct third-person pronoun, but the demonstratives can also function as third-person pronouns. In Yukulta (Tangkic; Queensland, Australia), for example, *ṉiya* is the third-person pronoun, but the demonstratives, *ṭanta* 'this one' and *ṭaṯinta* 'that one', can also function as third-person pronouns (Keen 1983: 213).

The remaining languages that belong to this group show a derivational relationship between third-person pronouns and demonstratives. This involves the use of the same stem in the case of seven languages, and the use of the third-person pronoun itself for deriving demonstrative pronouns (or vice versa) in the case of eight languages.

In Axininca Campa (Arawakan; central Peru), for example, the third-person pronouns *irirori* 'he' and *iroori* 'she' on the one hand, and the demonstratives *irika* (M), *iroka* (F) 'proximal', *irinta* (M), *ironta* (F) 'medial', and *irintó* (M), *irontó* (F) 'distal' on the other, all share the stem *ir-* (Reed and Payne 1986: 324, 330).

In Ambulas (Sepik; Papua New Guinea), on the other hand, demonstrative pronouns are derived by attaching the deictic markers *kén* 'proximal' and *wan* 'remote' to third-person pronouns (Wilson 1980: 53, 55). This is also true of Khasi (Mon-Khmer; northeastern India), in which the demonstratives are derived by adding the third-person pronouns *u* (M) and *ka* (F) to *-ne* 'proximal', *-to* 'medial', and *-tay* 'remote' (Nagaraja 1985: 10–11).

Kharia (Munda; central India) shows the reverse type of relationship in that the third-person pronouns are derived by attaching *kar* 'person' to the demonstratives *u* 'this', *ho* 'that', *han* 'remote', and *hin* 'invisible' (Biligiri 1965: 37).

1.3 Third-person pronouns related to remote demonstratives. The affinity between third-person pronouns and demonstratives is restricted to remote demonstratives in the case of eighteen languages. Again, the affinity can consist in exact identity or in a shared stem. In the case of nine of these eighteen languages, third-person pronouns are identical with remote demonstratives. In Eastern Armenian, for example, the third-person pronoun is *na*, which is identical to the remote demonstrative *na*, but distinct from the proximal demonstrative *sa* and medial demonstrative *da* (Kozintseva 1995: 12–13).

On the other hand, there are nine languages in which the third-person pronouns share their stem with the remote demonstratives. For example, in Lower Grand Valley Dani (Trans-New Guinea; southern Papua, Indonesia), *at* 'third person' shares a stem with the remote demonstrative *aty* 'that', which is distinct from the proximal demonstrative *jy* 'this' (Bromley 1981: 190, 207).

Five of these nine languages show an affinity of this type only between the nonsingular form of the third-person pronoun and the remote demonstrative. For example, Hawaiian (Polynesian) has the third-person pronoun *ia*, which assumes the stem *lā* in the dual and plural, and the distal demonstrative *kēlā* (which has the alternants *lā* and *ana*) (Elbert and Pukui 1979: 107, 110).

1.4 Third-person pronouns related to non-remote demonstratives. The affinity is restricted to non-remote demonstratives in the case of fourteen languages. The affinity is with proximal demonstratives in the case of four of these languages. In Asmat (Trans-New Guinea; southern Papua, Indonesia), for example, the third-person pronoun is *ar*, which also functions as the proximal ('near 1') demonstrative, contrasting with the two nonproximal pronouns *ja* 'near 2' and *er* 'distal' (Voorhoeve 1965b: 142, 155).

In the case of the remaining ten languages, the affinity is between third-person pronouns and medial demonstratives. Brahui (Dravidian; Pakistan), for example, has the demonstratives *dā* 'proximal', *o* 'medial', and *ē* 'distal', of which the medial one also functions as third-person pronoun (Bray 1986: 84).

In the case of eight of these languages, the demonstrative that is related to the third-person pronoun is not specified as "medial", but it does nevertheless contrast with a distinct 'distal (yonder)' demonstrative. For example, Lezgian (Nakh-Daghestanian; eastern Caucasus) has the demonstratives *im* 'proximal', *am* 'distal', and *at'a* 'remote (yonder)', of which *am* functions as third-person pronoun (Haspelmath 1993: 184–8).

1.5 Same gender markers. In the case of 24 languages, the affinity derives from the recurrence of the same set of gender (or noun class) markers. For example, Apurinã (Arawakan; northwestern Brazil) has the third-person pronouns *u-wa* 'he' and *o-wa* 'she', and demonstrative pronouns *u-kira* (M), *o-kira* (F) 'distal' and *i-ye* (M), *o-ye* (F) 'proximal' (Facundes 2000: 350, 361). Similarly, the third-person pronouns of Venda (Bantu; South Africa) show a distinction for 21 noun classes, and each of the language's four demonstrative pronouns also makes a similar distinction for the 21 noun classes (Poulos 1990: 95, 106).

Nine of these languages are described as showing a masculine-feminine gender distinction, and three as showing a masculine-

feminine-neuter gender distinction; twelve of them, on the other hand, are described as showing different kinds of noun class distinctions. In the case of one of these languages, Grebo (Kru; Liberia), the distinction is between large, important things (including humans) and small, unimportant things (Innes 1966: 46, 50).

In the case of some of the languages in which third-person pronouns are identical with demonstratives (values 2–4), we do find gender distinctions occurring among forms that function as both third-person pronouns and demonstratives. For example, Malayalam (Dravidian; Kerala, India) has demonstratives *avan* (M), *avaḷ* (F) 'remote', and *ivan* (M), *ivaḷ* (F) 'proximal' that can also function as third-person pronouns (Asher and Kumari 1997: 258), and these show a gender distinction. However, such languages are not included in this fifth group. Only languages in which third-person pronouns are distinct from demonstratives, but are nonetheless similar in sharing the same set of gender markers, are included in this group.

1.6 Demonstratives used as nonhuman third-person pronouns. In the case of the remaining seventeen languages, the affinity derives from the fact that third-person pronouns are used only for referring to human or animate beings. In order to refer pronominally to other beings or inanimate objects, these languages make use of their demonstrative pronouns. For example, the third-person pronoun *upa* of Jaqaru (Aymaran; Peru) is used for referring to humans only; the demonstratives *aka* 'this' and *uka* 'that' are used for referring to nonhumans (Hardman 2000: 27).

2 Problems of classification

2.1 Degrees of derivational relationship. Since the notion of affinity or "derivational relationship" is a graded one, the division between the first group (unrelated) and the rest (related) cannot be sharp and clear-cut. It is quite possible that some of the languages that are included in the first group could also be classified as belonging to one of the other groups. For example, Serbian-Croatian has the third-person pronouns *ôn* 'he', *ònа* 'she', and *ònо* 'it', and the demonstrative pronouns *ovaj* 'near 1', *taj* 'near 2', and *onaj* 'distal' (Kordić 1997: 22–5). It is possible that one or more of these pronouns (especially *onaj* 'distal') share the stem with the third-person pronouns. Similarly, Mundari (Munda; northwestern India) has the third-person pronoun *ay*, which appears to resemble the demonstrative pronouns *niy* 'proximal', *iniy* 'medial', and *hiniy* 'remote' in its ending (Osada 1992: 64, 68). Eastern Pomo (Hokan; northern California) has the pronouns *mi·p'* 'he' and *mi·ṭ'* 'she', which appear to be correlatable with the proximal demonstrative *mé* 'this' (McLendon 1975: 124, 161). However, none of these languages is regarded here as showing any affinity between third-person pronouns and demonstratives.

Third-person pronouns are considered to be related to demonstratives only in the case of languages in which either (i) the affinity is quite obvious or (ii) the author has suggested that the two are historically related.

2.2 Absence of independent third-person pronouns. Two of the languages of the sample are reported to have no independent third-person pronouns as such. They are Cayuvava (isolate; Bolivia) and Wichita (Caddoan; Oklahoma, Kansas, Texas). In Cayuvava, words that translate as pronouns are derived by attaching the direct object personal affixes to certain auxiliary stems. These also function as auxiliary verbs (Key 1967: 34). In Wichita, on the other hand, personally inflected participles of the verb 'be' are given as translations of English pronouns (Rood 1976: 10).

However, in Cayuvava, the verbal prefixes for denoting the person and number of the subject show a deictic distinction in the case of the third person (*/ϕ/ ~ /ir/ ~ /t/ ~ /r/* 'he, she, it' versus */k/* 'he, she, it (far away)'). Words that translate as pronouns are derived by attaching object suffixes (rather than subject prefixes) to the auxiliary verb *ãre* 'is, exists' and therefore do not show this deictic distinction; but we do find such a distinction in the case of the forms *me arenahi* 'this one' and *me arekehi* 'that one' (Key 1967: 28, 36). I have therefore included this language under group two. On the other hand, Wichita has the demonstratives *ti?i* 'this', *ha:ri* 'that', the latter of which can occur as a third-person pronoun (Rood 1976: 10). I have therefore included this language under group three.

3 Geographical distribution

These typological distinctions do not appear to manifest any striking patterns in their geographical distribution. It might perhaps be suggested that languages showing no relation (group one) are more common (i) in Africa, (ii) in South-East Asia stretching north to Japan and Korea and east into the Pacific, and (iii) in North America, except for California, than they are in other parts of the world.

4 Two-person and three-person languages

The question as to whether third-person pronouns are related or unrelated to demonstratives appears to form the basis for a typological distinction between two distinct sets of languages. We might call languages in which third-person pronouns are related to demonstratives **two-person languages**, and those in which the two are unrelated **three-person languages**. Third-person pronouns form part of the system of personal pronouns in the latter case but not in the former. This distinction appears to be correlatable with certain other characteristics such as, for example, that

(i) gender distinctions among third-person pronouns are displayed more frequently by two-person languages than by three-person languages, and

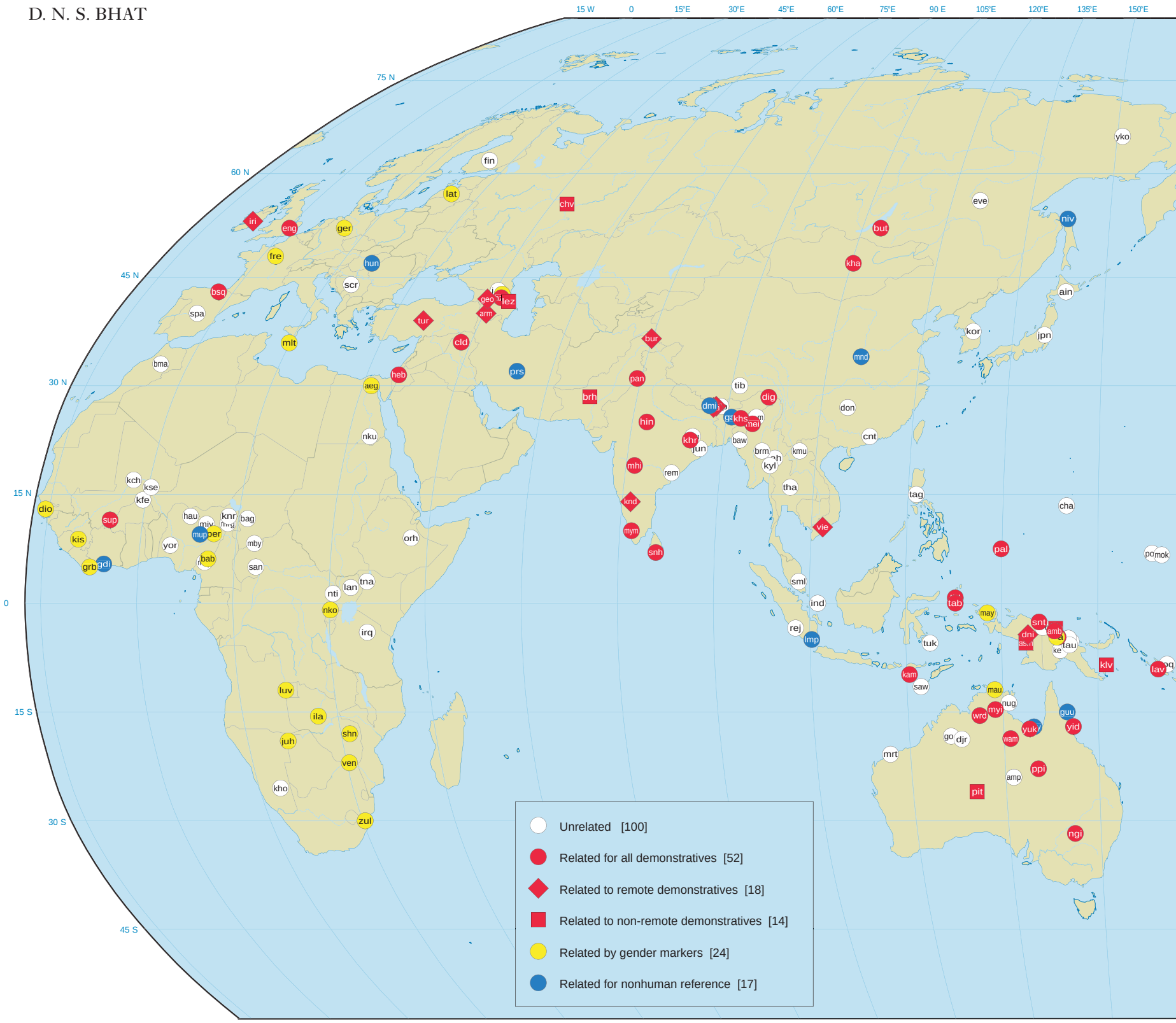
(ii) the deictic system is less frequently person oriented in two-person languages than in three-person languages.

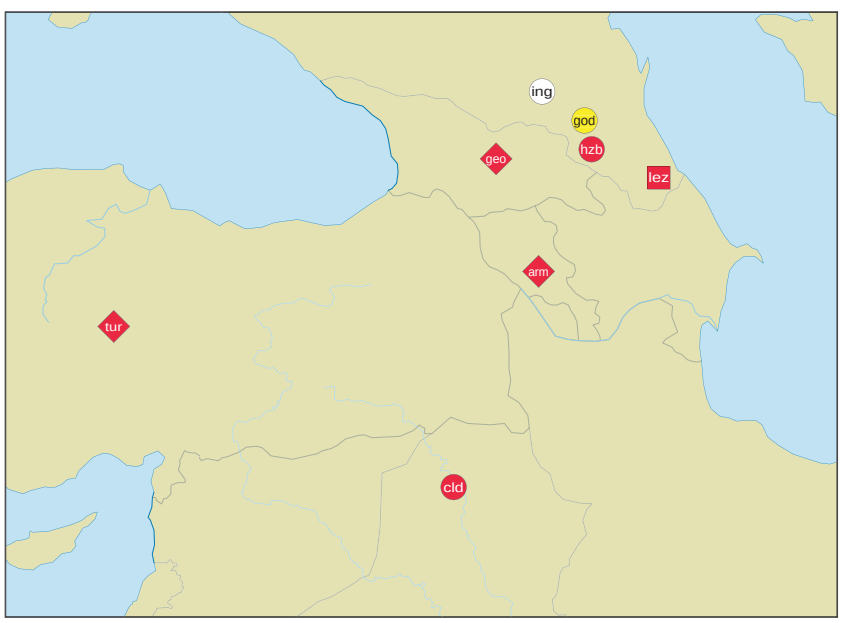
For example, among the 125 two-person languages that occur in my sample, 43 show a gender distinction, whereas among the 100 three-person languages, only 13 show a gender distinction (among seven of these latter languages, the distinction occurs in second person as well). This disparity strongly supports the first correlation mentioned above.

There is a similar kind of disparity in the occurrence of person-oriented deictic systems (involving 'near the addressee' as a distinct entity) which supports the second correlation given above. Among the 125 two-person languages that occur in the sample, only fourteen have a person-oriented deictic system, whereas among the 100 three-person languages, as many as 31 have a person-oriented deictic system (see Bhat 2004).

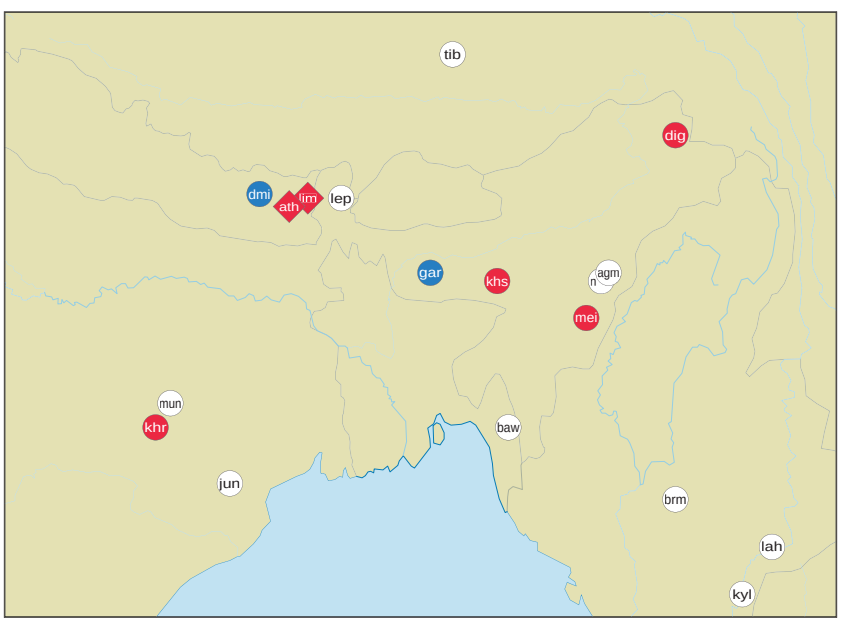
43 Third-Person Pronouns and Demonstratives

D. N. S. BHAT

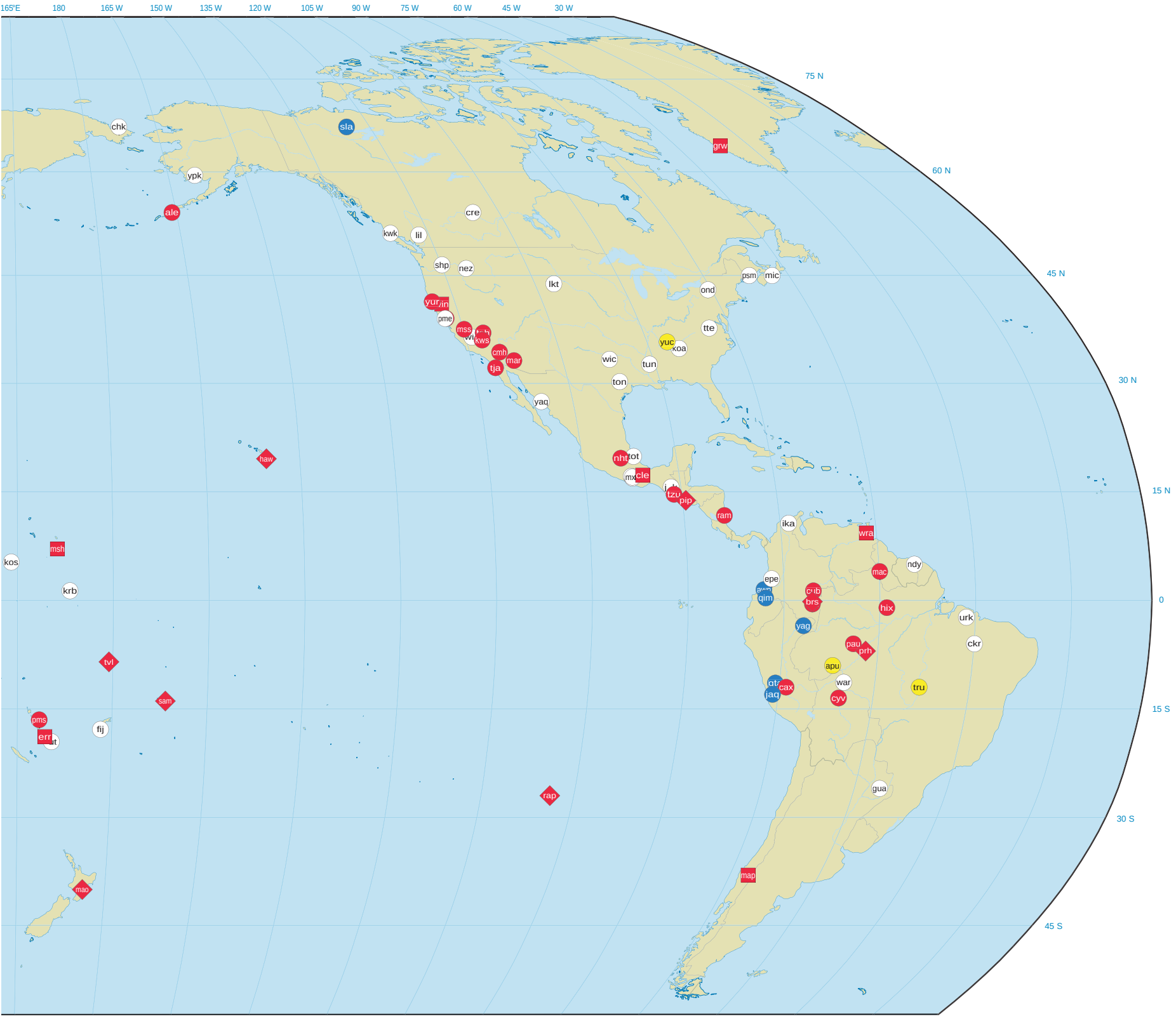


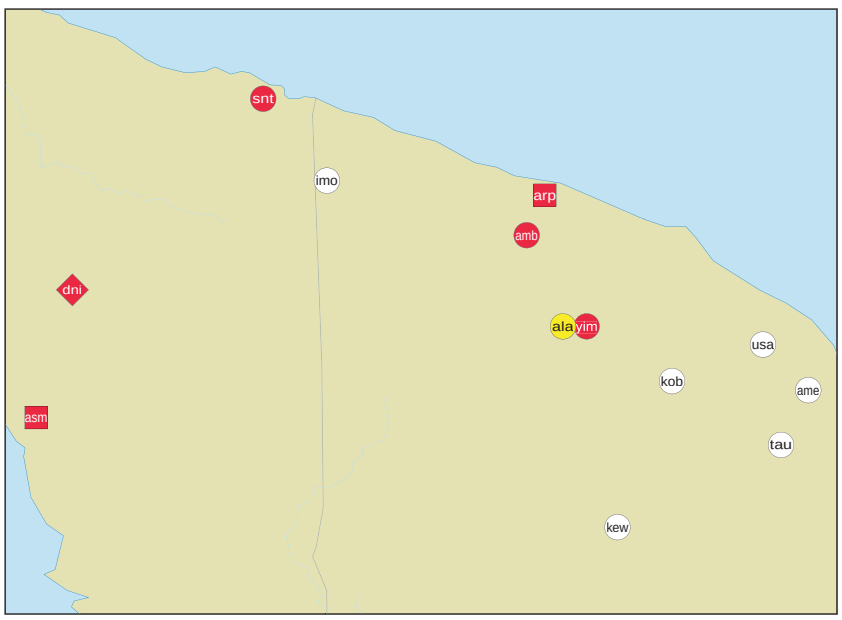


Caucasus and adjacent areas

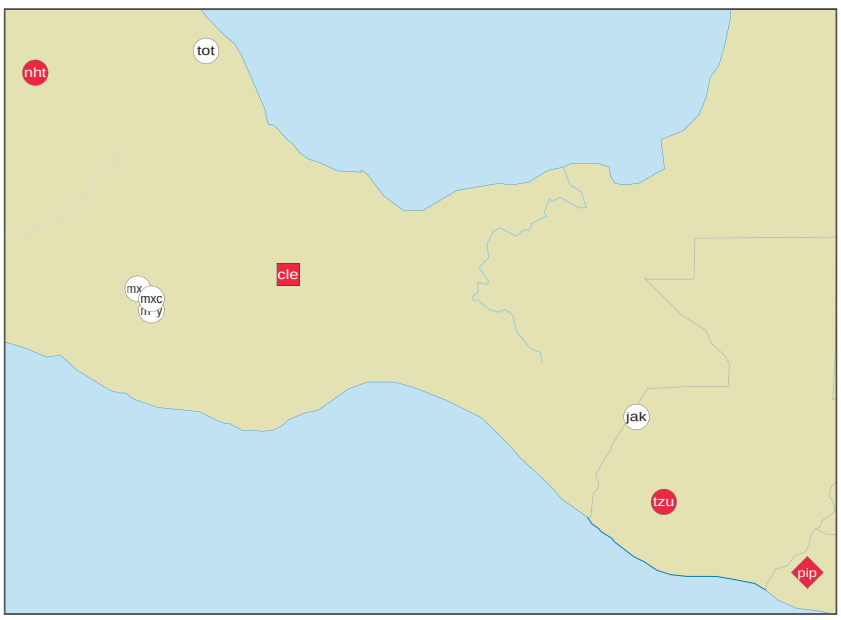


Nepal, eastern India, and surrounding areas





central New Guinea



Mesoamerica

44 Gender Distinctions in Independent Personal Pronouns

ANNA SIEWIERSKA

1 Introduction

Independent personal pronouns, in contrast to bound forms (i.e. clitics and affixes), are separate words capable of taking primary stress. Virtually all languages have independent personal pronouns, though in some they occur rather infrequently. Gender, as defined by Corbett (1991), is a form of classification of nominals, as shown by agreement (see Map 30). Most gender contrasts on personal pronouns are sex-based, i.e. pronouns used for male referents are masculine and those used for females are feminine. The treatment of other referents varies. They may be referred to by a separate set (or sets) of neuter pronouns, as is the case in English and many other European languages. Alternatively, they may be grouped with the referents of masculine gender (e.g. Amharic), or less commonly with the referents of feminine gender (e.g. Warekena), or split over the masculine and feminine genders in an arbitrary way or according to some semantically based principle (e.g. Garifuna).

Gender oppositions in personal pronouns are characteristic of the third rather than the first or second person. This is suggested by Greenberg's (1963: 96) Universal 44: "If a language has gender distinctions in the first person, it always has gender distinctions in the second or third person or in both." Moreover, gender is seen as being typical of singular rather than nonsingular personal pronouns. This, in turn, is expressed in Greenberg's Universal 45: "If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also." There are some exceptions to Universals 44 and 45 involving independent personal pronouns, but not many (cf. especially Plank and Schellinger 1997).

2 Defining the values

The distribution of gender marking on independent personal pronouns is shown in Map 44. In the case of languages which have more than one series of independent personal pronouns, the forms considered are those that may be used as subjects. The six values depicted in the map are based on the relationship between gender and person as reflected in the above-mentioned Greenbergian universals. The values are:

1.	Gender distinctions in third person plus first and/or second person	18
2.	Gender distinctions in third person only, but in both singular and nonsingular	42
3.	Gender distinctions in third person singular only	61
4.	Gender distinctions in first or second person but not third	2
5.	Gender distinctions in third person nonsingular only	1
6.	No gender distinctions	254
	total	378

The first value represents languages in which gender is highly prominent in the pronominal system, i.e. in addition to gender in the third person, there is also gender in either second person or first person or both. The gender distinctions in question may involve just the singular, or any combination of both singular and nonsingular. For instance, in Hausa (Chadic; Nigeria and Niger) and many other Chadic and Semitic languages, gender is exhibited only in the second and third person singular:

(1) Hausa (Newman 2000: 477)

1SG	*nī*	1PL	*mū*
2SG.M	*kai*	2PL	*kū*
2SG.F	*kē*		
3SG.M	*shī*	3PL	*sū*
3SG.F	*ita*		

In Ngala (Ndu, Sepik; northwestern Papua New Guinea), there are gender distinctions in all three persons in the singular:

(2) Ngala (Laycock 1965: 133)

1SG.M	*ɯn*	1DU	*ʌyn*	1PL	*nan*
1SG.F	*ñən*	2DU	*ən*	2PL	*gɯn*
2SG.M	*mən*	3DU	*kəbər*	3PL	*rʌr*
2SG.F	*yn*				
3SG.M	*kər*				
3SG.F	*yn*				

Rif Berber (Morocco) has a gender contrast in the second and third person in both the singular and plural (see (3)), while the Murui dialect of Huitoto (Huitotoan; Colombia and Peru) has a two-way gender contrast in the first and second person dual and a three-way gender contrast in the third person singular, dual, and plural (see (4)).

(3) Rif Berber (McClelland 2000: 27)

1SG	*nəš*	1PL	*nəšnin*
2SG.M	*šək*	2PL.M	*kəniw*
2SG.F	*šəm*	2PL.F	*kənint*
3SG.M	*nətta*	3PL.M	*nitnin*
3SG.F	*nəttæθ*	3PL.F	*nitənti*

(4) Murui Huitoto (Wise 1999: 322)

SG		DU		PL	
1	*kuè*	1 M	*kɔkɔ*	1	*kaũ*
		1.F	*kaɨ̀naù*		
2	*ɔ*	2 M	*ɔmukɔ*	2	*ɔmɔù*
		2.F	*ɔmuɲɔù*		
3.M	*imue*	3 M	*iaùmaìaù*	3.M	*imakɯ*
3.F	*iɲaìɲɔ*	3.F	*iaùɲuaù*	3.F	*iɲaìɲuaù*
3.N	*ie*	3 N	*ie*	3.N	*ie*

And Korana (Central Khoisan; South Africa) has gender distinctions for all three persons in three numbers and in both the exclusive and inclusive forms in the dual and plural (here C = common gender):

(5) Korana (Meinhof 1930: 43, via Tom Güldemann)

		SG	DU	PL
1	C		*sa-m*	*sa-da*
INCL	F		*sa-sam*	*sa-sē*
	M		*sa-kham*	*sa-tjē*
1	C		*si-m*	*si-da*
EXCL	F	*ti-ta*	*si-sam*	*si-sē*
	M	*ti-re*	*si-kham*	*si-tjē*
2	C		*sa-khao*	*sa-du*
	F	*sa-s*	*sa-saro*	*sa-sao*
	M	*sa-ts*	*sa-kharo*	*sa-kao*
3	C	*‖'ãi-'i*	*‖'ãi-kha*	*‖'ãina*
	F	*‖'ãi-s*	*‖'ãi-sara*	*‖'ãidē*
	M	*‖'ãi-b*	*‖'ãi-khara*	*‖'ãiku*

The Korana paradigm is particularly worthy of attention as it is the most complete person/number/gender paradigm in the sample.

Languages in which gender in personal pronouns is confined to the third person but not just to the singular are covered by value

two. The relevant languages may display gender in all the singular and nonsingular third-person forms, as in Worora (Wororan; Western Australia), which has gender distinctions in the singular, dual, trial, and plural.

(6) Worora (Love 2000: 8–10)

	SG	DU	TRI	PL
1INCL		*ŋa'rendu*	*'ŋariŋ'guri*	*'ŋari*
1EXCL	*'ŋaiu*	*a'rendu*	*'ariŋguri*	*'ạri*
2	*'ŋundju*	*nji'rendu*	*'njiriŋguri*	*'njiri*
3.M	*'indja*	*iŋ'gandu*	*'iŋguri*	*'arka*
3.F	*'njina*	*njiŋ'gandinja*	*'njiŋgurinya*	*'arka*
3.N1	*'ẉuna*	*ẉun'gạndu*	*'ẉunguri*	*'ẉuna*
3.N2	*'mạna*	*mạn'gạndum*	*'mạngurim*	*'mạna*

Alternatively, they may have gender in the singular and only some of the nonsingular numbers, for instance, the dual but not the plural, as in Lavukaleve (Solomons East Papuan; Solomon Islands):

(7) Lavukaleve (Terrill 2003: 170, 172)

	SG	DU	PL
1EXCL	*ngai*	*el*	*e*
1INCL		*mel*	*me*
2	*inu*	*imil*	*imi*
3.M	*fona*	*fonala*	*fova*
3.F	*fo*	*fol*	*fova*
3.N	*foga*	*fogala*	*fova*

(In Lavukaleve four degrees of distance are distinguished in the third-person forms. Only the proximal forms are given in (7).)

Value three encompasses languages which have a gender contrast solely in the third person singular, irrespective of the variety of number oppositions exhibited: singular/plural as in English, singular/dual/plural as in Mundari (Munda; India), Trumai (isolate; Brazil), and Yessan-Mayo (Sepik; Papua New Guinea), or even singular/dual/paucal/plural as in Ungarinjin (Wororan; Western Australia).

Value four represents languages which display gender in either the first or second person but not the third. The only instance of gender just in the first person that I have come across with respect to independent personal pronouns is in Macá (Mataco-Guaicuru; Argentina and Bolivia), in which feminine gender is marked on the first person inclusive (Aikhenvald 2000: 252). Gender just in the second person is slightly more common. It is illustrated in (8) from Iraqw (Cushitic, Afro-Asiatic; Tanzania), which has both long and short forms of independent personal pronouns.

(8) Iraqw (Mous 1993: 112)

	long	short
1SG	*aníng*	*án*
2SG.M	*kúung*	*kú*
2SG.F	*kíing*	*ki*
3SG	*inós*	*ís*
1PL	*atén*	*át*
2PL	*kuungá*	–
3PL	*ino ín*	*inín*

Some other languages which have gender just in the second person are Burunge (Cushitic, Afro-Asiatic; Tanzania), Kofyar (Chadic, Afro-Asiatic; Nigeria), and Minangkabau (Sundic, Austronesian; Sumatra, Indonesia).

Value five has been assigned to languages which display gender only in the third person nonsingular. These languages constitute a subtype of exceptions to Greenberg's Universal 45. A case in point is Dagaare (Gur; Ghana), which has a human/nonhuman distinction only in the third person plural. This is illustrated in (9).

(9) Dagaare (Bodomo 1997: 71)

1SG	*maa*	1PL	*tenee*
2SG	*foo*	2PL	*yɛnee*
3SG	*onɔ*	3PL.HUM	*bana*
		3PL.NONHUM	*ana*

Plank and Schellinger (1997: 62–5) mention several other languages belonging to this type, which are not in the sample, namely the Sauias dialect of Biak, Wandamen, and Windesi (Austronesian; Papua, Indonesia); Katu (Katuic, Mon-Khmer; Vietnam and Laos); and Palauan (Western Malayo-Polynesian; Palau Island).

Finally, value six represents languages that have no overt gender contrasts in the independent personal pronouns. Included in this category are languages in which third-person pronouns are confined solely to humans or animates while other referents are referred to by demonstratives, classifiers, or full NPs. Such languages are sometimes interpreted as displaying a covert gender opposition.

3 Geographical distribution

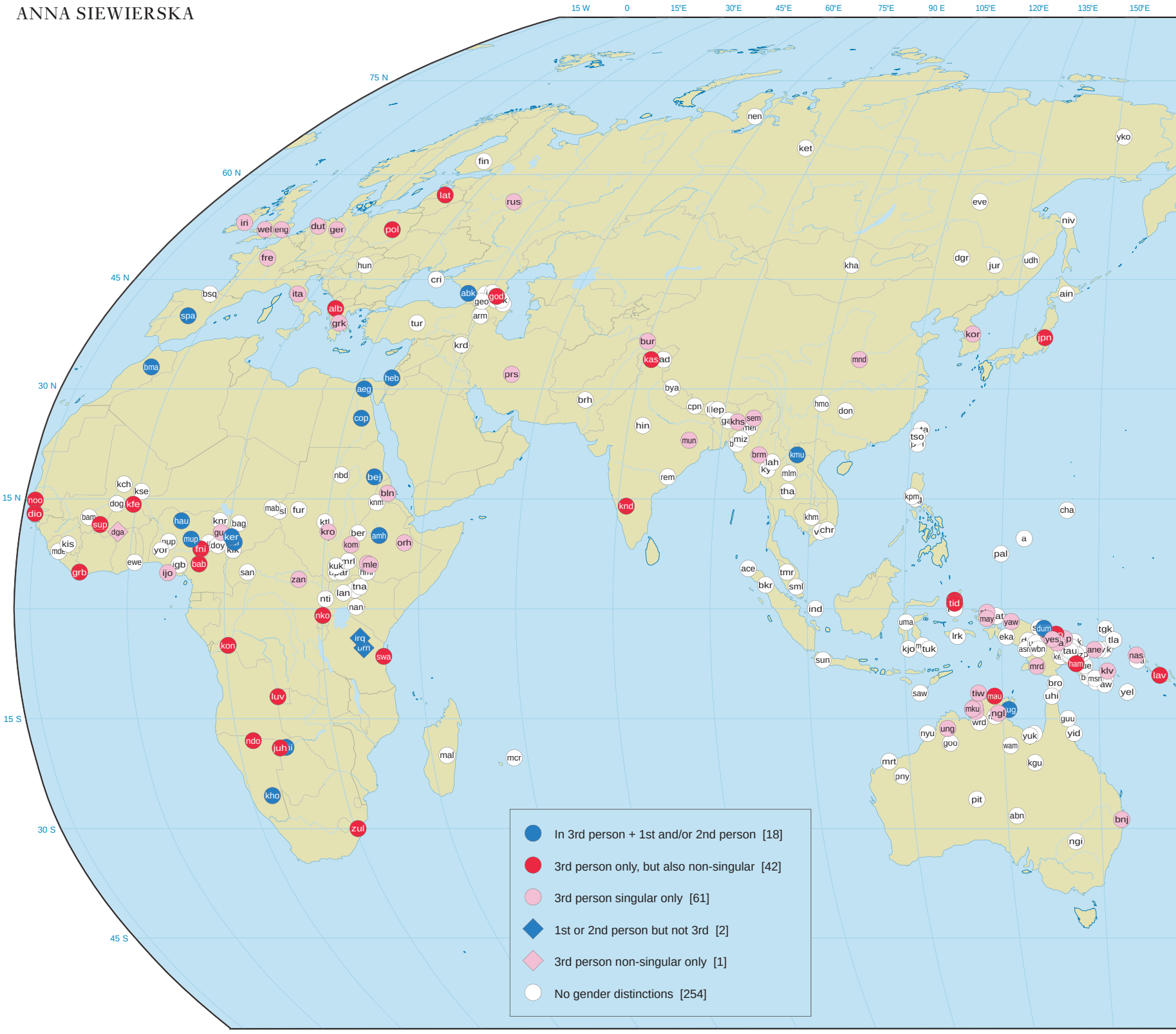
Gender in independent pronouns is found in about 30 per cent of the languages in the sample. The greatest concentration of languages with gender in personal pronouns is in Africa. Gender is very prominent among the Afro-Asiatic languages of northern Africa, in the Niger-Congo languages of sub-Saharan Africa, and also in the Khoisan languages of the southern part of the continent. It is not, however, a regular feature of Nilo-Saharan languages. The other major area with gender distinctions in independent pronouns is Eurasia and especially Europe. Gender is also very much in evidence among the non-Pama-Nyungan languages of northern Australia.

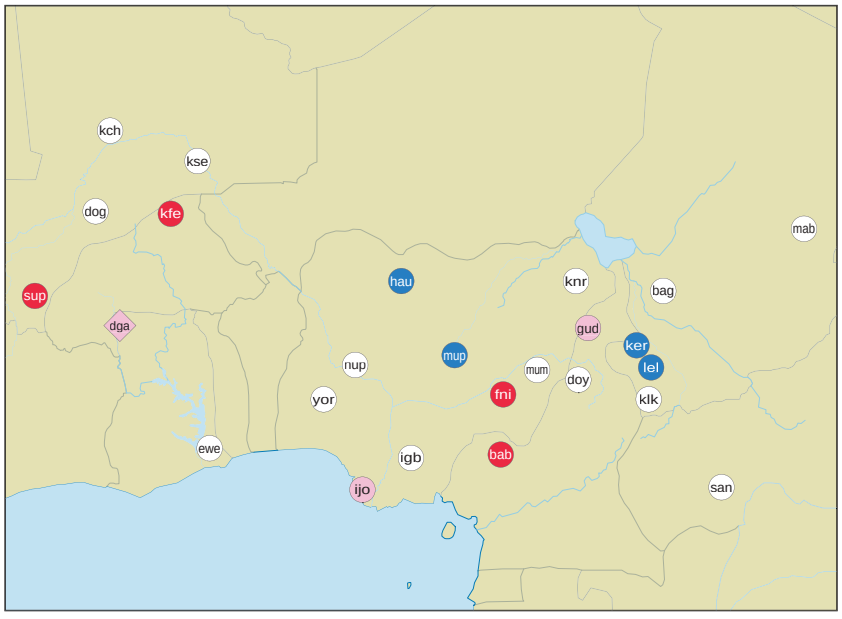
Turning to areas where gender is much less prominent, in New Guinea gender is found mainly among the non-Trans-New Guinea languages of northern New Guinea, particularly in the Sepik area and neighbouring areas of West Papua. It also occurs among some of the West Papuan languages on the extreme north-west tip of New Guinea and in northern Halmahera. There is even less gender in the Americas, especially North America. The North American languages that do have gender are scattered throughout the continent. In South America gender is found primarily in the northern part of the continent and especially among the languages of the Amazon basin. The area in which gender is encountered most rarely is South-East Asia and Oceania. Gender occurs in some Austro-Asiatic languages but hardly at all in Austronesian languages. All but one (Drehu) of the languages exhibiting gender from this area are non-Austronesian.

Among the languages that do have gender distinctions in independent pronouns, gender just in the third singular is dominant in all areas but Africa. In Africa the dominant pattern seems to be gender in third person irrespective of number. In the case of the Afro-Asiatic and Khoisan languages, the favoured distribution of gender is in both the second and third person. Outside of Africa gender in other than the third person, particularly in the singular, is rare. Instances of gender in the first and second person nonsingular are found in Europe (Spanish, Lithuanian, Slovene), New Guinea (Baniata, Dumo), and Australia (Anindilyakwa, Ndjébbana, Nunggubuyu). Only in Africa do we find languages displaying gender exclusively in some person other than the third.

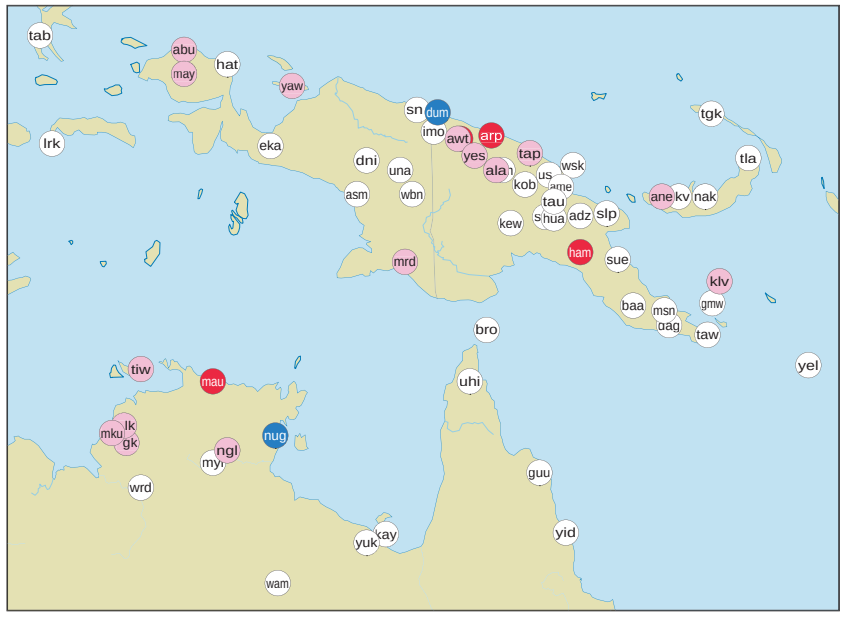
44 Gender Distinctions in Independent Personal Pronouns

ANNA SIEWIERSKA

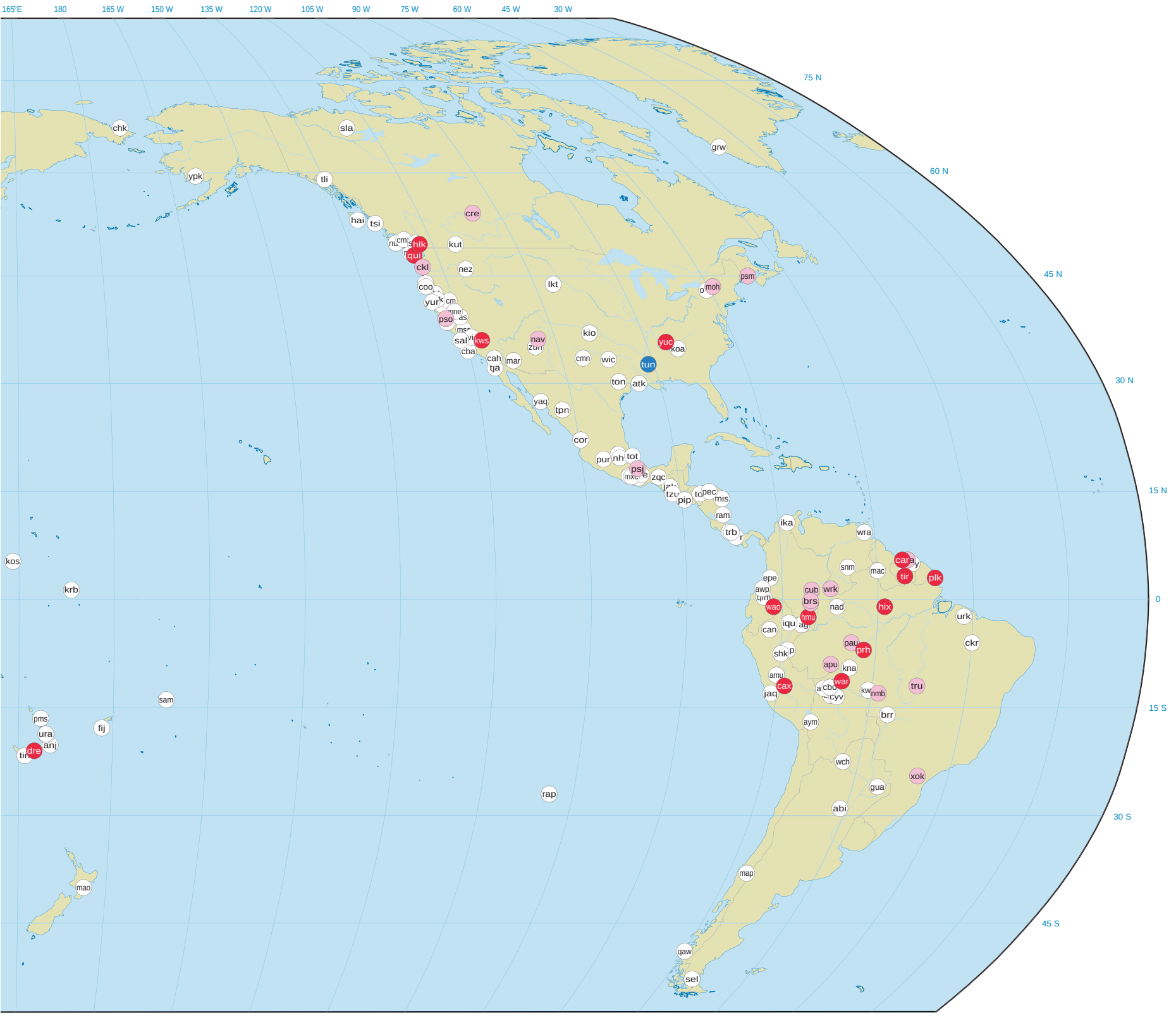


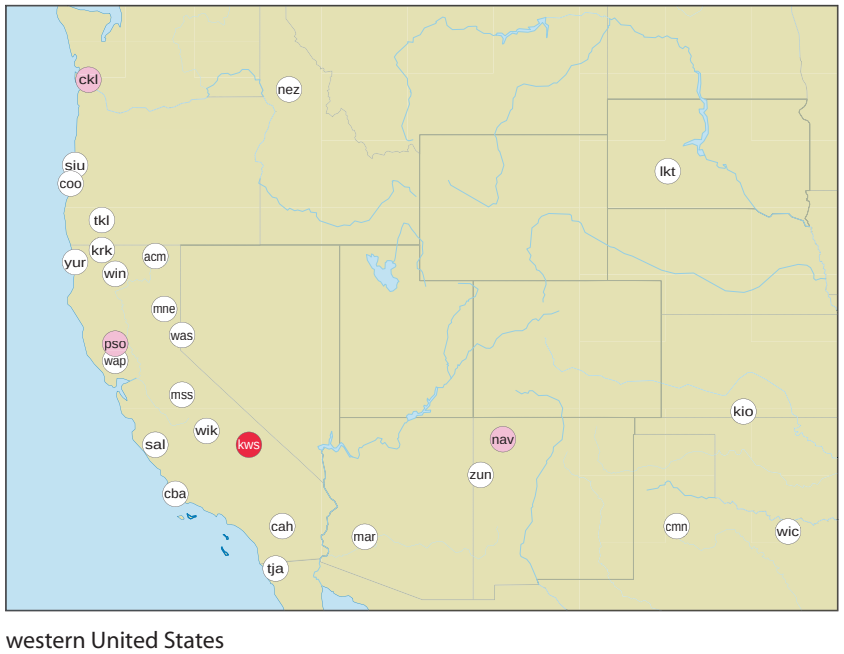


Nigeria and surrounding areas

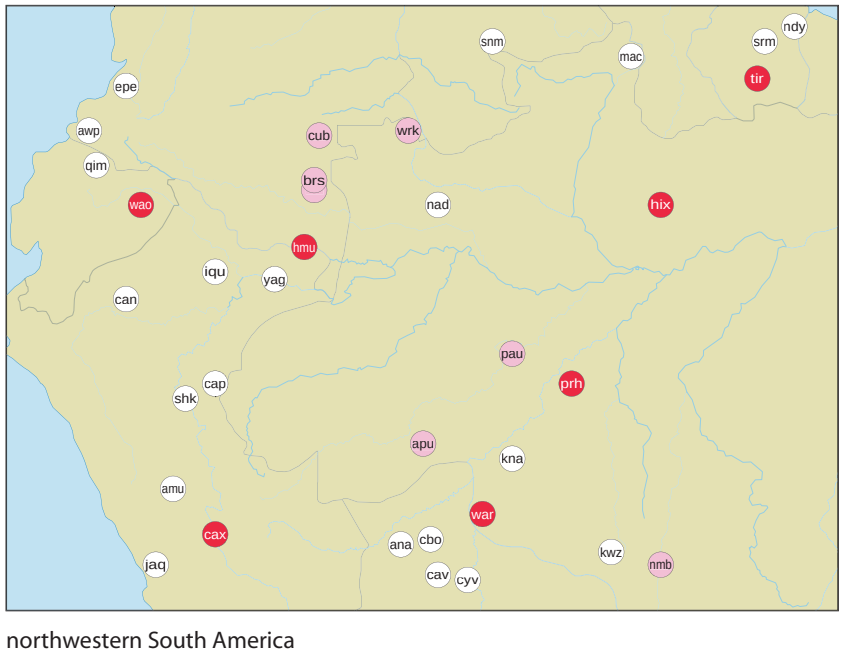


New Guinea and northern Australia





western United States



northwestern South America

45 Politeness Distinctions in Pronouns

JOHANNES HELMBRECHT

1 Politeness in pronouns

Map 45 is concerned with politeness distinctions in personal pronouns, and more specifically second-person pronouns. This phenomenon has been particularly well-studied in European languages. In German, for instance, there is a threefold distinction between a second person singular *du* 'you.SG.FAMILIAR', a second plural *ihr* 'you.PL.FAMILIAR', and a second person polite *Sie* 'you.HONORIFIC', which does not distinguish number. The polite pronoun *Sie* is the default form of address among German adults who are not in a close social relation to each other. If they are relatives, married, close friends, professional colleagues, or university students, they are very likely to use the familiar *du* to address each other. The usage of *du* and *Sie* is symmetrical (the one who gives *Sie* (or *du*) will also receive it), with one exception: Adults address children and adolescents with *du*, but receive *Sie*.

The simplified example of the usage of German *du* and *Sie* given here demonstrates that a single grammatical distinction (familiar versus polite) corresponds to a complex set of pragmatic rules and social contexts determining the choice between the two forms. In addition, the pragmatic rules which determine the choice of *du* and *Sie* are not identical to those of other languages with a binary politeness distinction in pronouns. Although there is a significant degree of overlap in the conditioning factors of the usage of *du* and *Sie* in German, *tu* (you.SG.FAM) and *vous* (you.HON) in French, and *ty* (you.SG.FAM) and *vy* (you.HON) in Russian, there are also significant pragmatic differences. In French, for example, the mother-in-law continues to be addressed by the (new) son-in-law or daughter-in-law with the polite form *vous*, whereas in German, the mother-in-law always receives the familiar form from the (new) son/daughter-in-law. In Russian, the usage of *ty* and *vy* is also conditioned by the affective relation between the interlocutors, the topic, and the discourse context (cf. Friedrich 1966, 1972). Interlocutors can change from the polite form of address to the familiar form and back again with certain pragmatic effects. This is not possible in a conversation of German interlocutors. Once two German interlocutors have agreed on the familiar form of address, they cannot return to the polite form.

2 Definition of politeness values

The scope of this chapter is restricted to politeness distinctions in second-person pronouns: politeness distinctions in first-person and third-person pronouns are excluded (for an overview of politeness distinctions in first- and third-person pronouns, see Shibatani 1998 and Helmbrecht 2004: chs. 8, 10). In addition, only referential expressions which can be classified as personal pronouns are considered. For the majority of the languages in the sample, this does not raise any problems, as the forms in question are part of a well-organized paradigm of structurally and distributionally similar forms. A problem arises, however, in some South-East Asian languages such as Thai, Burmese, and Vietnamese, where personal pronouns do not constitute an easily identifiable word class. Rather, there is a continuum stretching from pronominally used forms resembling the European type of personal pronouns, to forms which are more similar to nouns. In these cases, only those forms which are used exclusively for second-person reference are taken into consideration, regardless of whether they are also used as nouns. In Vietnamese, for instance, there is a large class of kinship nouns such as *anh* 'elder brother' and *ông* 'grandfather' (cf. Cooke 1968: 127–30) which are frequently used in pronominal function. But these can be used for first- and third-person reference, as well as for second-person reference. These nouns are therefore excluded from the database. By contrast, other forms which are etymologically nouns have acquired an exclusively second-person reference with a specific politeness value. Often, nominal and pronominal uses coexist. These forms are included in the database.

No difference was made between bound pronominal affixes and pronouns which are free independent words. Languages with bound pronouns indicating degrees of respect are found in Mesoamerica, e.g. in Nahuatl (Uto-Aztecan) and Mixtec (Oto-Manguean) languages.

The values for this map are indicated in the value box:

Second-person pronouns	
1. encode no politeness distinction	136
2. encode a binary politeness distinction	49
3. encode multiple politeness distinctions	15
4. are dominantly avoided for politeness reasons	7
total	207

The four honorific values for second-person pronouns are explained in the subsequent sections.

2.1 No politeness distinctions. This feature value is self-explanatory. Languages that were assigned this value have no personal pronouns in their paradigms which are used to express different degrees of respect or intimacy toward the addressee.

2.2 Binary politeness distinctions. This feature value covers all languages with the European type of politeness distinction (German *du*/*Sie*, French *tu*/*vous*, Russian *ty*/*vy*). These languages have a paradigmatic opposition between one intimate or familiar pronoun of address and another one expressing respectful address. Such a binary distinction may also be expressed by several distinct pronouns. It may be the case, for instance, that two different pronouns indicating the same degree of respect are used in different dialects. This is the case with *wy* and *Pan*/ *Pani* in Polish: the former is used in rural areas, the latter in urban areas (cf. Mühlhäusler and Harré 1990: 145–50). Or, different respect pronouns may be used in different grammatical contexts, e.g. one is a free pronoun fulfilling all grammatical functions, the other is an obligatory clitic only in subject function. This is the case with Taba (Austronesian; Indonesia) *meu* (2.SG.HON free pronoun) and *h*= (2.SG.HON clitic pronoun) (cf. Bowden 2001: 187–9). Criterial for this honorific type is that the pronouns do not indicate more than one politeness distinction.

2.3 Multiple politeness distinctions. This feature value covers all languages which exhibit two or more degrees of politeness within a pronominal paradigm. These systems are rare cross-linguistically. A typical example can be found in Marathi (Indo-Aryan; India). There is a form *tū* used for family members and intimate persons, two forms with the same degree of respect *te* and *he* (2SG.HON) for people with higher social status, and an extra polite form *āpaṇ* (2SG.HHON) for priests and teachers and in very formal contexts (cf. Pandharipande 1997: 375–94).

2.4 Pronoun avoidance. This feature value is the most difficult to identify of the four, and is terminologically rather different from the first three. The names of the first three values describe a categorical feature of the pronominal paradigms to be investigated. The term "pronoun avoidance", however, describes a strategy of pronoun usage which has an effect on the overall shape of the paradigm. Languages of East and South-East Asia such as Japanese, Burmese, and Thai have a strong sensitivity to politeness in language usage

and within their grammars. Speakers have to account for a variety of social distinctions linguistically. Social distinctions between speaker and hearer may reflect relative age, kinship, social ranking, intimacy, and other social features. From a linguistic point of view, one of the most important strategies of being polite is to avoid addressing people directly. This strategy of negative politeness (cf. Brown and Levinson 1987: 129–210) is the functional background of the emergence of politeness distinctions in pronouns in European languages, and indeed holds in languages both with and without grammaticalized politeness distinctions. The peculiarity of the languages of (South)-East Asia, however, is that personal pronouns are not used in polite address at all — instead, status and kinship terms, titles, and other complex nominal expressions are employed. The effect of this strategy on the shape of pronominal paradigms in these languages is that there are rarely polite pronouns of address. If there are second-person pronouns they are used to address social equals and inferiors. Polite forms of address, e.g. for the address of superiors, mostly do not belong to the class of personal pronouns in these languages. Since the discrimination of a separate class of personal pronouns is not as clear-cut as in European languages, e.g. in terms of binary paradigmatic oppositions, I prefer to specify this feature value in terms of the overall strategy, instead of using a term which presupposes paradigmatic generalizations for these languages which would necessarily be vague and difficult to measure.

3 Frequencies and areal distribution

The cross-linguistic survey of politeness distinctions in pronouns reveals that this is not a marginal grammatical feature in the world's languages. Around three quarters of the languages have no politeness distinctions in personal pronouns. But of the remaining quarter which do have politeness distinctions in second-person pronouns, two thirds have a binary politeness distinction, about 10 per cent have multiple politeness distinctions, and around 20 per cent must be classified as languages which express politeness by means of pronoun avoidance. These figures are based on a more representative subset of languages than is the actual 207-language sample used. The sample of 207 languages is somewhat biased towards European and (South)-East Asian languages, where politeness distinctions in pronouns are a widespread phenomenon.

The geographical distribution of the different honorific types across the languages of the world is not even. As the map clearly indicates, there are large areas where politeness distinctions in pronouns are totally absent, and there are areas where this phenomenon occurs with a remarkable density. Areas where politeness is not a category in personal pronouns are North and South America, New Guinea, Australia, and most of Africa.

Pronominal paradigms with a binary politeness distinction (value two) occur everywhere else but show a hotbed in Europe and adjacent areas.

Multiple politeness distinctions (value three) in pronouns are rare, but occur predominantly in languages of South Asia and neighbouring areas. This feature could be considered a defining feature of this area. Many of the languages there use a second-person plural pronoun for the first degree of politeness, and a third-person or reflexive pronoun for the expression of the next higher degree of politeness. The second-person singular form is usually reserved for close equals, family members, and inferiors.

Pronoun avoidance (value four) is a defining feature of the South-East Asian area. Languages such as Thai, Burmese, Vietnamese, and others often lack any polite second-person pronoun, but have many forms for reference to equals, inferiors, or for impolite usage. These lexical gaps with respect to polite second-person pronouns are characteristic of the paradigms in these languages.

4 Linguistic politeness: theoretical issues

A very influential attempt to provide a general account of the usage of second-person polite pronouns is Brown and Gilman (1960). The basic idea is that the usage of familiar/polite second-person pronouns (in European languages) is largely determined by two not fully independent macro-sociological parameters, power and solidarity. Polite pronouns may be used for address when there is a difference in social rank and prestige between the interlocutors; the asymmetrical use of second-person pronouns reflects the difference in the social power of the interlocutors. This parameter seems to be relevant in German only with respect to the usage of *du* and *Sie* between adults and children. The solidarity parameter pertains to the social distance of the interlocutors. If they are strangers there is a greater social distance between two interlocutors than if they are members of the same social group, such as family, school, professional organization, and so forth. Polite pronouns are used if the social distance is great, familiar counterparts are used if the social distance is less. The usage of German *du/Sie* largely reflects differences along this scale. Solidarity-based pronoun usage is always symmetrical (for a critical overview of this approach, see Agha 1994).

A theory which incorporated the two parameters of Brown and Gilman, but is much wider in its scope, is Brown and Levinson (1987). Their theory tries to give a general explanation of (linguistic) politeness based on the social-psychological notion of face. Face is described as the "public self-image that every member wants to claim for himself" (Brown and Levinson 1987: 61). This public self-image, face, is a part of the personality of each individual, and has to do with the way the individual wants to be seen and treated by others in the society. Brown and Levinson's basic idea is that there are numerous types of speech acts and utterances which threaten the face wants of the addressee.

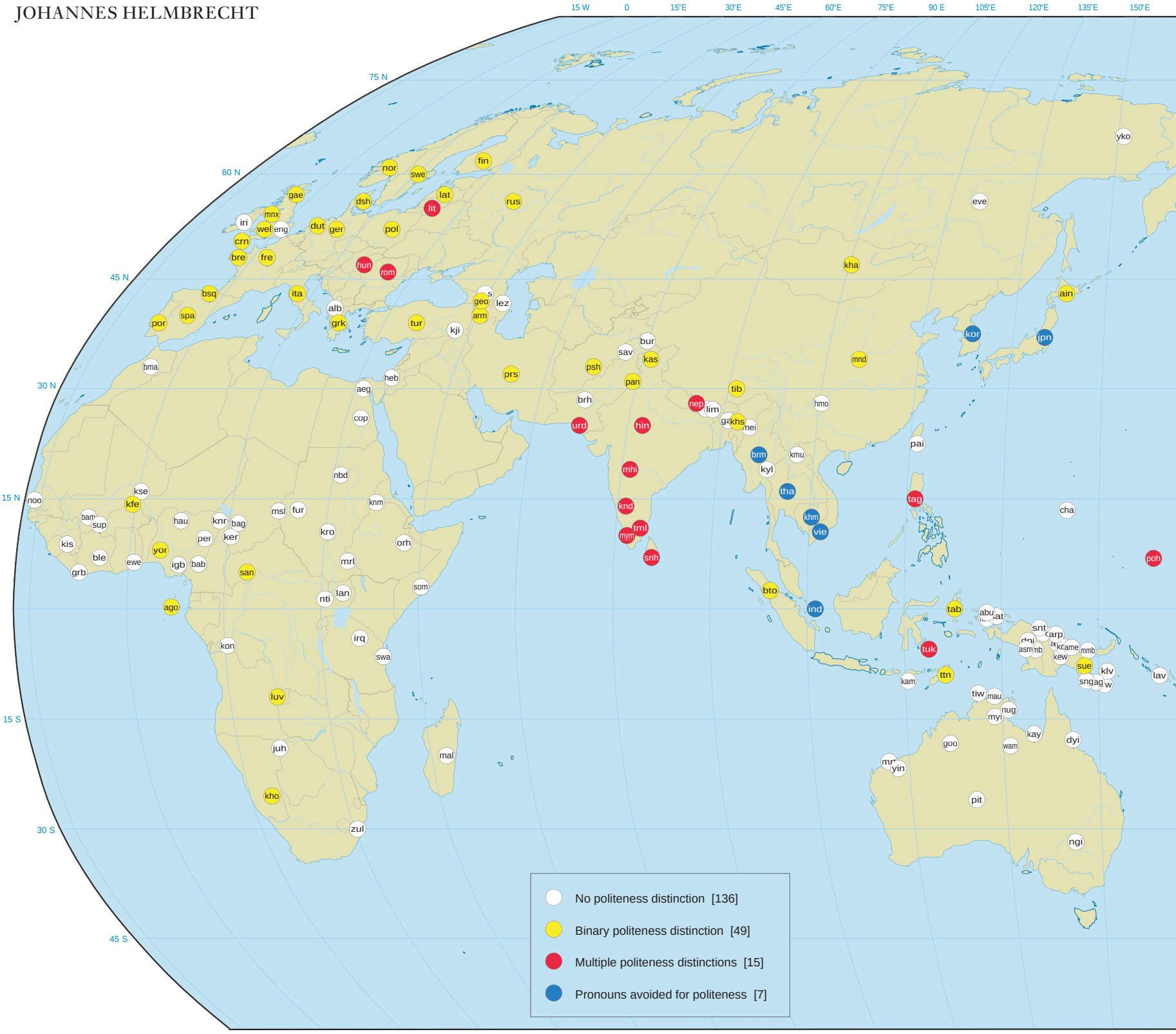
The avoidance of direct (linguistic) reference to the addressee in the context of face-threatening utterances is the main functional motivation for developing polite referential expressions such as *vous* in French and *Sie* in German. The 2PL pronoun *vous* in French presumably came into use historically as a polite form of singular address because it renders the reference less direct and less specific (cf. Malsch 1987, Helmbrecht 2003, 2004). Other possible diachronic sources for second-person polite pronouns are first-person plural pronouns (for example in Ainu (Japan)), demonstrative pronouns (for example in Sinhala (Indo-Aryan; Sri Lanka)), reflexive pronouns (for example in Hungarian), and nouns and nominal expressions designating social status (as in Spanish). All these sources of polite second-person pronouns avoid a direct second-person reference in the sense that they initially required some pragmatic inferencing before they were conventionalized as a polite means for pronominal reference. For a more detailed treatment of the functional aspects of the grammaticalization of second person polite pronouns, see Helmbrecht (2004: ch. 9).

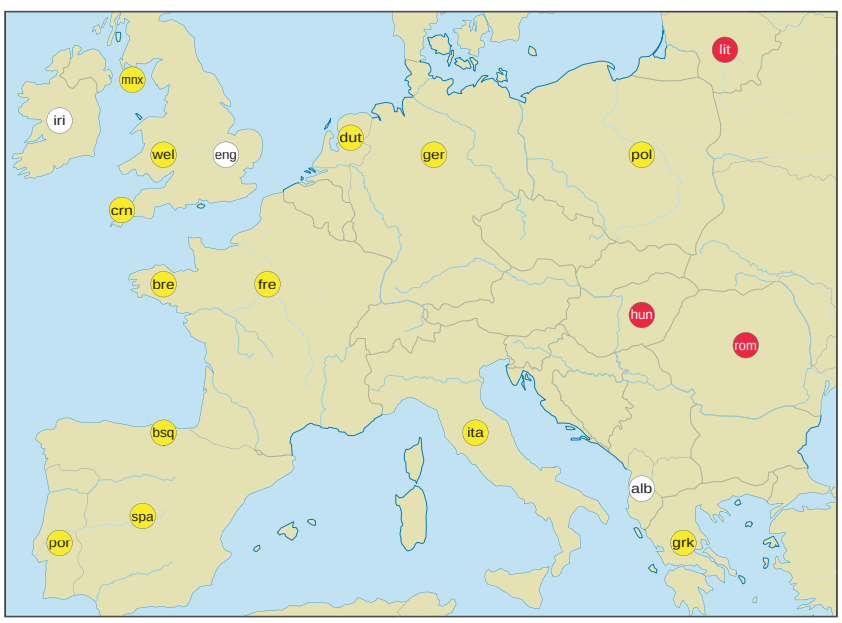
5 Conclusions

The uneven distribution of politeness distinctions in pronouns across the languages of the world suggests that there are other conditioning factors that have to be taken into account. Language contact and the social and cultural disposition to adopt linguistic means which are used to express politeness in neighbouring languages that have a high prestige seem to be more important as a determining factor than the general functional background of polite language use. It is this social and cultural disposition of the adopting society which is responsible for the selection of certain forms as politeness forms.

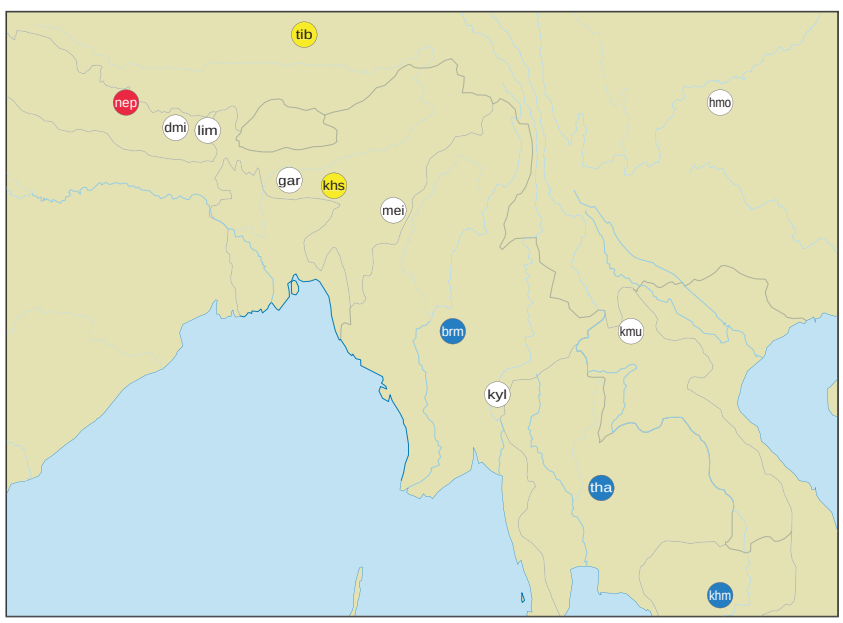
45 Politeness Distinctions in Pronouns

JOHANNES HELMBRECHT

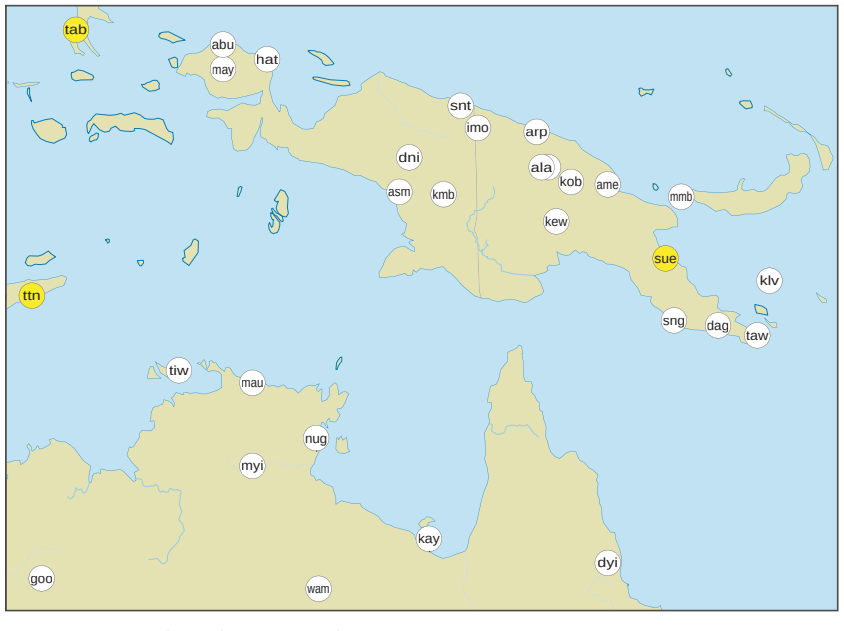




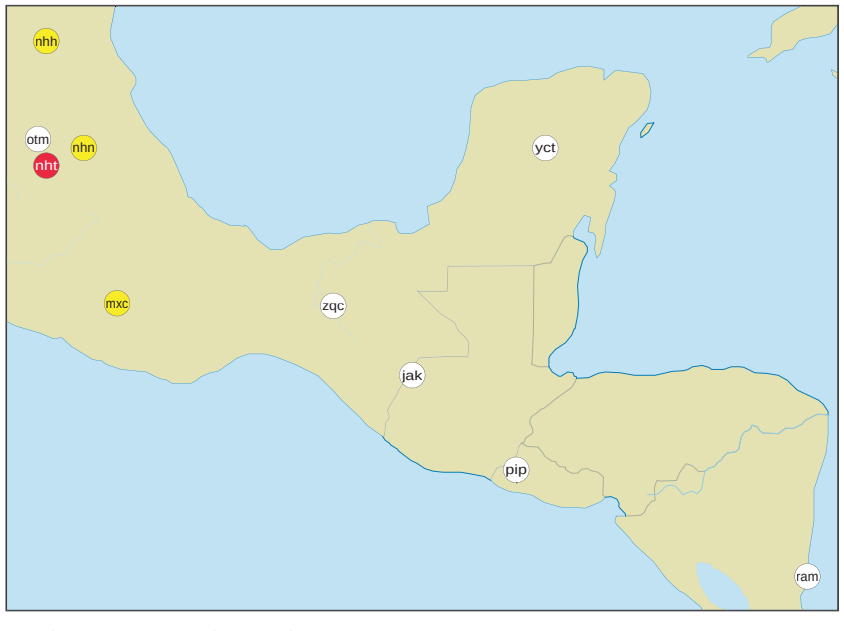
Europe



from Nepal to Cambodia



New Guinea and northern Australia



southern Mexico and Central America

46 Indefinite Pronouns

MARTIN HASPELMATH

1 Overview of the values

This chapter examines **indefinite pronouns** like 'somebody' and 'something' and distinguishes two major types. On the one hand, they may be closely related to the **interrogative pronouns** 'who' and 'what', as in Russian, where *kto-to* 'somebody' and *čto-to* 'something' are derived from *kto* 'who' and *čto* 'what' by means of the indefinite suffix *-to*. On the other hand, they may be closely related to the **generic nouns** 'person' and 'thing', as in Persian, where 'somebody' is expressed as *kæs-i* [person-INDEF] and 'something' is expressed as *čiz-i* [thing-INDEF]. In addition to the two major types, there are some languages with **special expressions** for 'somebody' and 'something', unrelated to anything else. An example is Abkhaz (North-West Caucasian), where 'somebody' is *aj^{ə}ə* and 'something' is *ak'ə* (Hewitt 1979: 158). Theoretically one might expect further types, e.g. indefinites derived from demonstrative pronouns, but no other types have been found.

Languages may also show a **mixed type**, e.g. when 'somebody' behaves differently from 'something'. For instance, in Khmer the word for 'somebody' is *kèː*, which is also translated as 'person' and is unrelated to *nɛək na:* 'who', but the word for 'something' is *ʔvɤy(-mú:əy)*, which is closely related to *ʔvɤy* 'what' (Jacob 1968: 134–5, 294, 322, 325).

Finally, languages may lack nominal indefinite pronouns entirely and express the equivalent of 'someone' and 'something' by means of an **existential construction**.

Thus, Map 46 shows five different values (see Haspelmath 1997: 242–3 for an earlier version of this map).

1.	Interrogative-based indefinites	194
2.	Generic-noun-based indefinites	85
3.	Special indefinites	22
4.	Mixed indefinites	23
5.	Existential construction	2
	total	326

In addition to the nominal indefinites 'somebody' and 'something', languages usually also have adverbial and adjectival indefinite expressions ('somewhere', 'somehow', 'for some reason', 'of some kind') and occasionally we even find verbal indefinites ('do something'). Typically, these behave in the same way as the nominal indefinites. Thus, in Nivkh (Russian Far East), indefinites of all kinds are derived from interrogatives, and in French, indefinites of most kinds are derived from generic nouns.

(1) Nivkh (Panfilov 1962: 253–7)

siḍ	'what'	*siḍ-lu*	'something'
aŋ	'who'	*aŋ-lu*	'somebody'
řain	'where'	*řain-lu*	'somewhere'
řaḍ	'which'	*řaḍ-lu*	'some'

(2) French

un	'one'	*quelqu'un*	'somebody'
chose	'thing'	*quelque chose*	'something'
part	'part'	*quelque part*	'somewhere'

Since many descriptions do not provide information on adverbial indefinites, this chapter is restricted to the nominal indefinites 'somebody' and 'something'.

2 The four types

2.1 Interrogative-based indefinites. In the majority of cases, interrogative-based indefinites are literally "based on" interrogative pronouns in the sense that they are overtly derived from them by the addition of a bound marker. A few examples are given in (3).

(3) a. Lak (Nakh-Daghestanian; Caucasus; Žirkov 1955: 74–5)
cu-riw 'somebody' *cu* 'who'

b. Modern Greek
ká-pjos 'somebody' *pjos* 'who'

c. Korku (Munda; India; Nagaraja 1999: 44)
je-ka 'somebody' *je* 'who'

d. Southeastern Tepehuan (Uto-Aztecan; Willett 1991: 55)
jaroi' 'somebody' *jaró* 'who'

This bound marker is typically related to expressions for 'be', 'want', 'perhaps', 'or', or 'also' (see Haspelmath 1997: chs. 6–7 for detailed discussion and diachronic explanations).

In other languages, indefinite and interrogative pronouns are identical. This possibility exists in some Slavic and Germanic languages (e.g. colloquial German *Ist da wer?* [is there who] 'Is there somebody there?'), and it is particularly widespread in Australian languages, e.g. Jaminjung (Jaminjungan; Northern Territory, Australia) *nanggayin* 'who; somebody' (Schultze-Berndt 2000: 45). It has been suggested that in these languages, indefinites and interrogatives are really identical and have the general meaning of "lack of information"; the specific interpretations as indefinite and interrogative arise from the larger construction or from the context (see Bhat 2000, 2004, as well as Mushin 1995 for Australian languages). If this is correct, these indefinites are not synchronically "interrogative based", but it appears that at least diachronically, the interrogative use is always primary (Haspelmath 1997: 176).

2.2 Generic-noun-based indefinites. A few languages clearly have dedicated indefinite pronouns which are based on the generic nouns 'person' and 'thing'. Instead of the noun 'person', some languages make use of the numeral 'one', and not uncommonly, these are accompanied by an indefinite determiner of some kind. A case in point is Italian, which has *qualc-uno* 'somebody' and *qual-cosa* 'something' (from *qualche* 'some' plus *uno* 'one' or *cosa* 'thing'), and English is similar (*some-one/some-body, some-thing*), although the noun *body* does not have the meaning 'person' anymore when it is used by itself. For these languages, we have to say that there are dedicated indefinite pronouns, because not all properties of the indefinites are predictable from their parts and from general rules (for instance the fact that *something* is stressed on the first syllable, as opposed to *some thing*).

But for most languages with generic-noun-based indefinites, there is no good evidence available that these expressions are different from ordinary indefinite noun phrases. In fact, some descriptions explicitly deny that they are indefinite pronouns. For Kobon (Trans-New Guinea), Davies (1981: 156) says: "There are no specific indefinite pronouns. This is expressed by an indefinite noun phrase such as *bɨ ap* 'a man', 'someone', *nan ap* 'something'." In this chapter, such cases where we lack evidence for pronominal status have been lumped together with languages like Italian and English because it would have been very difficult to draw a line between the two types. The evidence for pronominal status can be rather subtle. For instance, French *quelque chose* 'something' at first glance looks like a regular indefinite noun phrase consisting of the indefinite determiner *quelque* and the noun *chose* 'thing'. However, the notion 'something good' is expressed with a construction that is reserved for pronouns: *quelque chose de bon* (cf. *quoi de bon* 'what good'), whereas an ordinary noun phrase would be *quelque bonne chose* ('some good thing'). Unless a language has been described exceptionally thoroughly, such additional evidence is usually unavailable, but I suspect that it will eventually be found in more languages.

Thus, "pronominal status" was considered irrelevant for the purposes of this chapter, and indefinites were identified on purely semantic grounds: if a pair of expressions serves to render the notions 'somebody' and 'something', it is taken into account for this chapter, and if this pair contains the words 'person' (or 'one') and 'thing', then these indefinites count as generic-noun-based. Another possibility is that the indefinite determiner alone is used for 'somebody', as in Hausa (*wani/wata* 'some, someone', *wani àbù* 'something', *àbù* 'thing'; Newman 2000: 153–4).

Since many grammars of languages with generic-noun-based indefinites are simply silent about ways of expressing 'someone' or 'something', information often had to be extracted from dictionaries, example sentences, and texts. Even so, the map is likely to show some reporting bias: the actual proportion of generic-noun-based languages is likely to be higher than shown here.

2.3 Special indefinites. Special indefinites are indefinite pronouns whose roots are synchronically unrelated to other roots in the language. In a number of Indo-European languages, their diachronic origin is clearly in interrogative-based indefinites, but they have lost their synchronic connection to interrogatives. Examples are Spanish *alguien* 'somebody', from Latin *ali-quem* [INDEF-who], and *algo* 'something', from *ali-quod* [INDEF-what]; Latin *ali-* was a clearly segmentable prefix, but Spanish *alguien* and *algo* are synchronically monomorphemic. Many of the Indo-European languages of South Asia are like Spanish in this regard.

2.4 Mixed indefinites. This type is not particularly interesting because it subsumes diverse subtypes: the 'somebody' word may be generic-noun-based while the 'something' word is interrogative-based, as in Semelai (Mon-Khmer; Malay Peninsula): *sma?* 'person, someone', *mandemɔh* 'what, something' (Kruspe 1999: §6.2), or it may be the other way round, as in Maori: *wai raanei* 'somebody', *wai* 'who', *mea* 'thing, something' (Bauer 1993: 372). One of the two indefinites may be a special expression, e.g. Portuguese *alguem* 'somebody', not transparently related to *quem* 'who', versus *alguma coisa* [some thing] 'something'. Or a language may have more than one set of indefinites, e.g. German: one set is interrogative-based (*irgend-wer* 'someone', *irgend-was* 'something') and another set is special (*jemand* 'somebody', not transparently related to *Mann* 'man', and *etwas* 'something', not transparently related to *was* 'what').

Finally, in a few languages the same morpheme may have all three senses: generic noun, indefinite, interrogative. Thus, in Urubú-Kaapor (Tupi-Guarani; northeastern Brazil), *awa* is 'person, someone, who', and *ma'e* is 'thing, something, what' (Kakumasu 1986: 331, 342, 380). These languages were also classified as mixed.

2.5 ***Existential construction***. A language may completely lack nominal expressions for 'somebody' and 'something' and resort to an existential circumlocution instead. One of the two examples in my sample is Tagalog (Austronesian; Philippines), where 'somebody' and 'something' are expressed by an existential construction consisting of an existential particle plus a verb form:

(4) Tagalog (Schachter and Otanes 1972: 276)

May	*d<um>ating*	*kahapon.*
exist	<ACTOR.VOICE>come.PFV	yesterday

'Someone came yesterday (lit. There exists (one who) came).'

3 Indefinite meanings

Most languages have different sets of indefinite pronouns for different meanings or contexts. The contrast in English between *some*-indefinites (*somebody, something, somewhere*, etc.) and *any*-indefinites (*anybody, anything, anywhere*, etc.) is fairly typical. The *some*-indefinites must be used in affirmative declarative sentences with realis modality (*Somebody left/*Anybody left*), and the *any*-indefinites are used to convey a free-choice meaning (*You can take anything you like*, different from *You can take something you like*) and can be used in questions and conditional clauses (*Has anybody left? If anything disappears, . . .*). Different languages divide up the space of indefinite meanings in different ways (Haspelmath 1997: chs. 3–4).

Important in the present context is the fact that different indefinite meanings do not behave alike with respect to the main parameter of variation discussed here. Indefinites that express free choice, and more generally indefinites that translate as *any*-indefinites in English, seem to have a greater tendency to be interrogative-based than indefinites that are used in affirmative declarative realis sentences (roughly corresponding to English *some*-indefinites). Thus, in Hausa, **realis indefinites** (as we may call them for short) are generic-noun-based, but free-choice indefinites are interrogative-based:

(5) Hausa (Newman 2000: 153–4, 622–3)

realis indefinites		free-choice indefinites	
wani/wata	'some, somebody'	*koo-waa*	'anybody'
		cf. *wàa*	'who'
wani àbù	'something'	*koo-mee*	'anything'
cf. *àbù*	'thing'	cf. *mèe*	'what'
wani wurii	'somewhere'	*koo-yàushè*	'anywhere'
cf. *wurii*	'place'	cf. *yàushè*	'where'

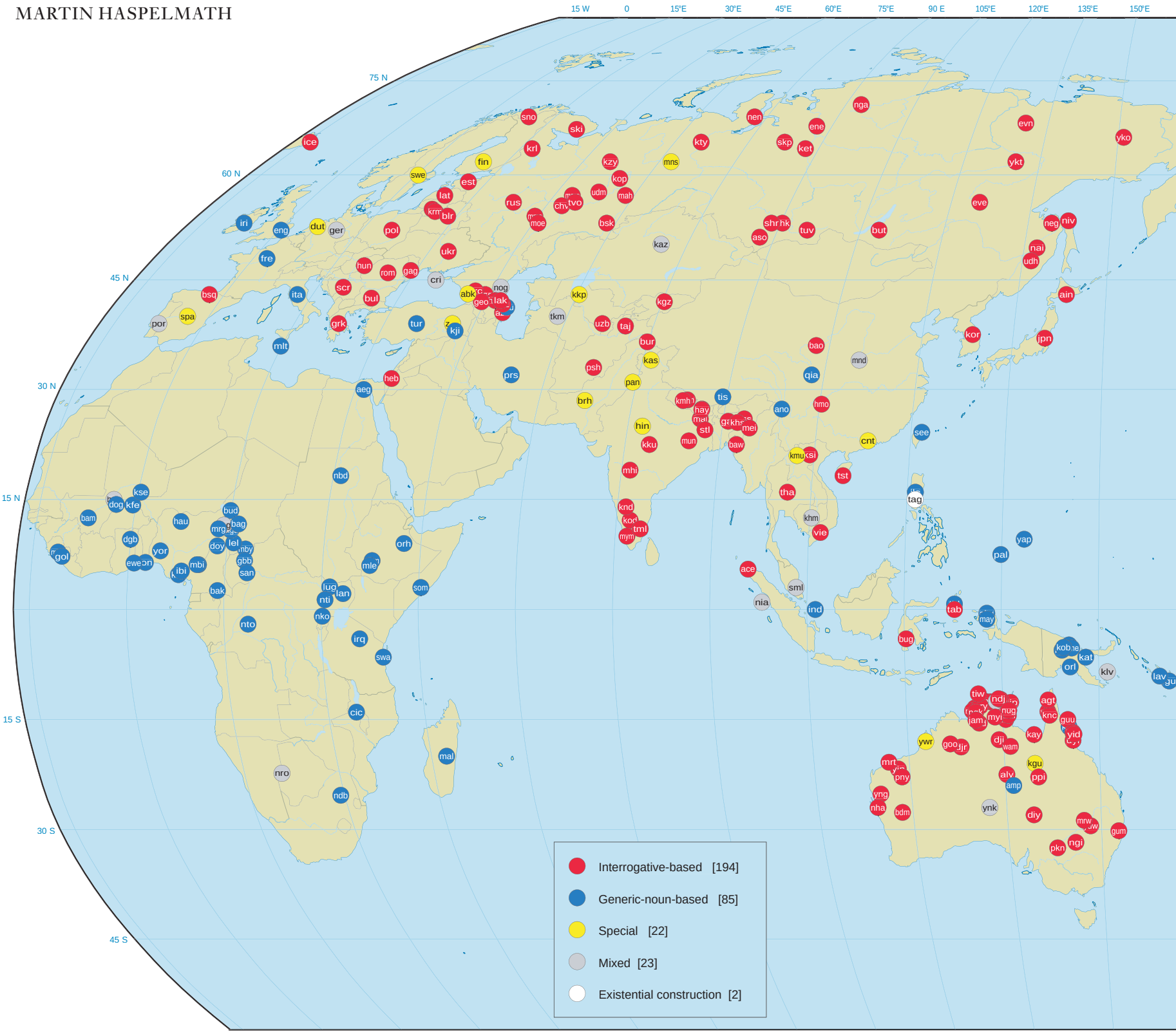
For the purposes of this chapter, only realis indefinites have been taken into account. In practice this means that indefinites translated as 'somebody/something' have been included, but indefinites translated as 'anybody/anything' have been disregarded.

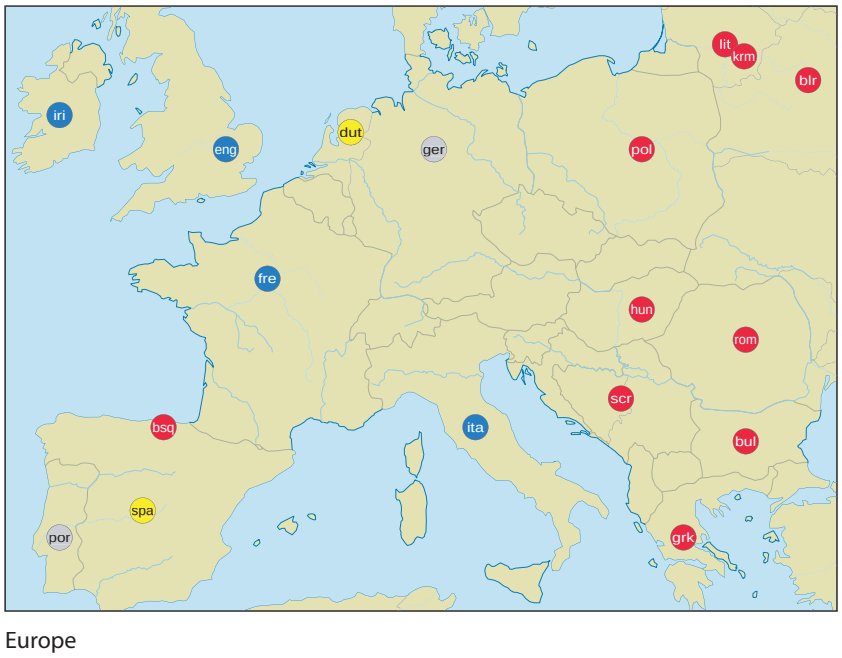
4 Geographical distribution

The geographical patterns in the distribution of the two main types are striking. Practically all African languages have generic-noun-based indefinites, and so do almost all the languages of New Guinea and the small Pacific islands. In striking contrast to this, the languages of North America, Australia, and Eurasia (apart from its western fringe) overwhelmingly show interrogative-based indefinites. The "special indefinites" type is scattered all over Eurasia, and the mixed type shows no clear pattern. This is one of the best examples of a feature whose distribution seems to be completely independent of language families and exhibits areal patterning on a continental scale.

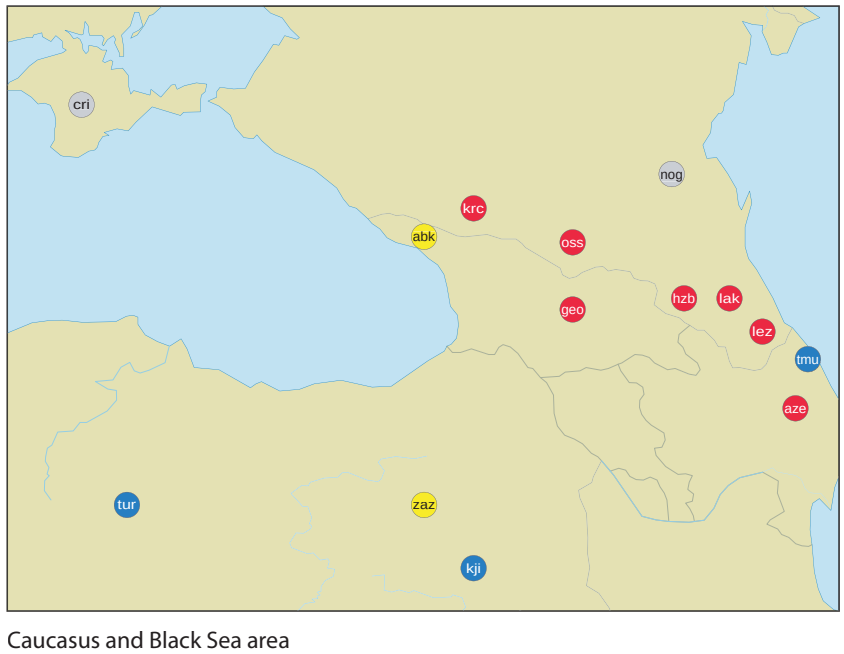
46 Indefinite Pronouns

MARTIN HASPELMATH

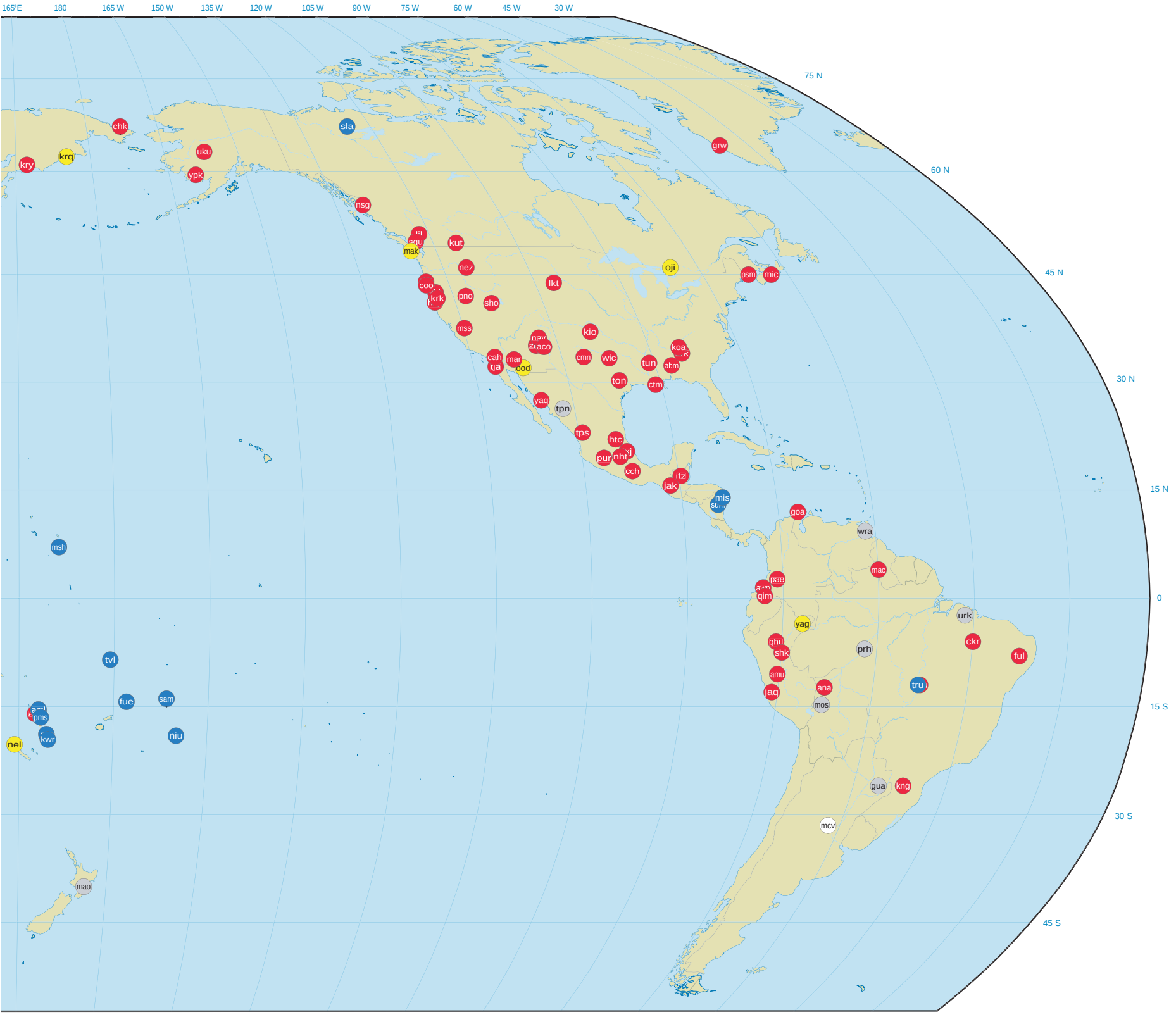


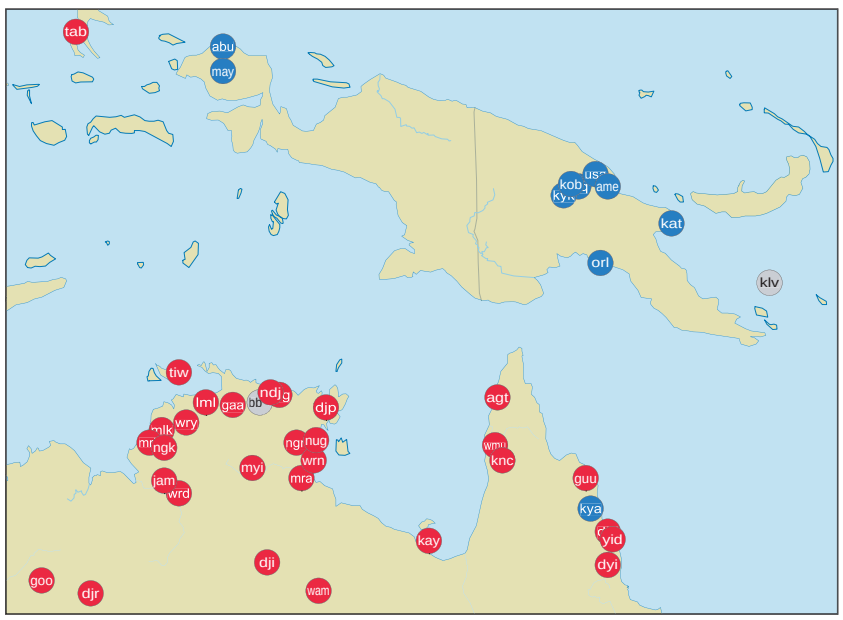


Europe

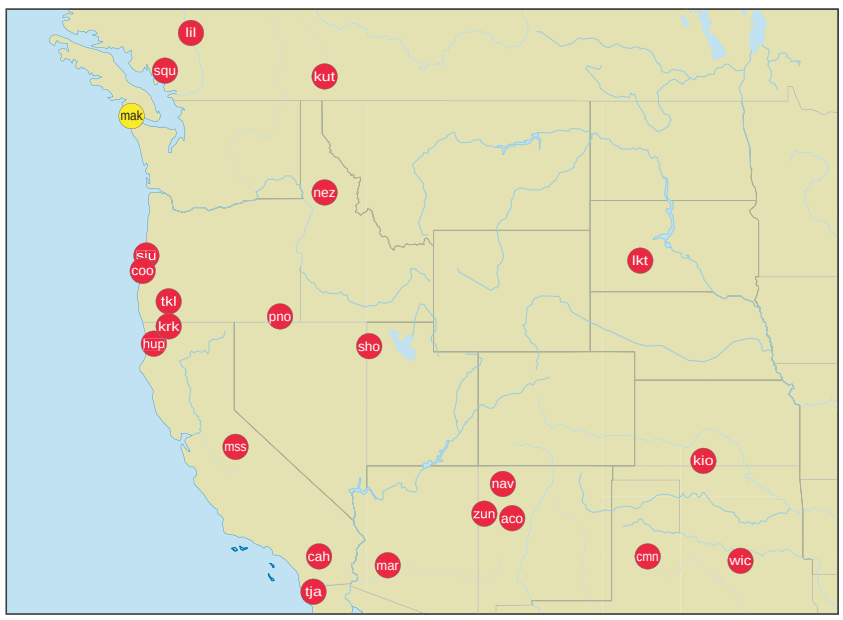


Caucasus and Black Sea area





New Guinea and northern Australia



western United States

47 Intensifiers and Reflexive Pronouns

EKKEHARD KÖNIG AND PETER SIEMUND
(with Stephan Töpper)

1 Defining the values

Reflexive pronouns (or "reflexive anaphors") are expressions which are prototypically used to indicate that a nonsubject argument of a transitive predicate is coreferential with (or bound by) the subject, i.e. expressions like German *sich*, Russian *sebja*, Turkish *kendi*, Mandarin *zìjǐ*, English x-*self*:

(1) *John*$_i$ *saw* ***himself***$_i$ *in the mirror.*

By **intensifiers** we mean expressions like German *selbst*, Russian *sam*, Turkish *kendi*, Mandarin *zìjǐ*, English x-*self*, which can be adjoined to either NPs or VPs, are invariably focused, and thus are prosodically prominent. The main function of intensifiers can be seen in the evoking of alternatives to the referent of the NP they relate to:

(2) a. (adnominal)
The director ***himself*** *opened the letter.*
b. (adverbial)
The director opened the letter ***himself****.*

Our main motivation for considering reflexives and intensifiers in tandem is that they are frequently **identical** in form and thus only differentiated in terms of distribution (as, for instance, in English). In languages in which they are formally **differentiated**, intensifiers can be used to reinforce reflexive pronouns. This is illustrated by (3c) from German, where the reflexive pronoun (*sich*, cf. (3a)) and the intensifier (*selbst*, cf. (3b)) are formally distinct.

(3) German

a.	*Johann*	*sah*	***sich***	*im Spiegel.*	
	Johann	saw	himself	in the mirror	
	'Johann saw himself in the mirror.'				
b.	*Der*	*Direktor*	***selbst***	*begrüßte*	*uns.*
	the	director	himself	welcomed	us
	'The director himself came to welcome us.'				
c.	*Paul*	*kritisierte*	***sich***	***selbst***.	
	Paul	criticized	himself	himself	
	'Paul criticized himsélf.'				

The map shows the geographical distribution of the two types of languages:

1.	Intensifiers and reflexive pronouns are formally identical	94
2.	Intensifiers and reflexive pronouns are formally differentiated	74
	total	168

2 Problems in assigning the values

The correct identification of intensifier-reflexive identity encounters the following three problems. First, there may be only **partial identity**. A typical situation, exemplified by Malayalam (Dravidian; southern India) in (4), is that the reflexive expression is a combination of intensifier and (personal or possessive) pronoun. For the purposes of the map, cases of partial and complete identity have been lumped together and contrasted with cases of nonidentity.

(4) Malayalam (Asher and Kumari 1997: 162)

a.	*avan*	***avane***	***taṉṉe***	*kuṟṟappeṭutti*
	he	him	INTF	accuse.PST
	'He accused himself.'			
b.	*ɲaan*	***taṉṉe***	*atə*	*ceyyaam*
	I	INTF	it	do.FUT
	'I myself will do it.'			

Secondly, languages may have **more than one reflexive pronoun**, only one of which may be (partially or completely) identical to the intensifier. A case in point is Dutch, where different verbs require different reflexive pronouns:

(5) Dutch
a. *Jan wast* ***zich****.*
'John washes.'
b. *Jan haat* ***zichzelf****.*
'John hates himself.'

Thirdly, there are languages that have **different intensifiers** for adnominal and for adverbial use, and that only use one of these (typically the adverbial one) as reflexive. This is illustrated by Japanese (cf. 6a-c; Ogawa 1998: 165–73)

(6) Japanese

a.	*Taro-wa*	***jibun-de***	*kuruma-wo*	*arrata.*	
	Taro-NOM	INTF-INSTR	car-ACC	washed	
	'Taro washed the car himself.'				
b.	*Taro*	***jishin***	*kyouju-wo*	*sonkeishiteiru.*	
	Taro	himself	professor-ACC	honour	
	'Taro himself honours the professor.'				
c.	*Taro-wa*	***jibun-wo***	*semeta.*		
	Taro-NOM	REFL-ACC	criticized		
	'Taro criticized himself.'				

In such problematic cases we have always chosen the interpretation that is maximally compatible with the idea of identity, i.e. Dutch and Japanese were counted as manifesting identity.

3 Geographical distribution

In the entire sample there are 61 languages that manifest complete identity of reflexive and intensifier. A further 33 languages show partial identity whereas 74 languages draw a clear formal distinction between the two expressions. The combination of cases of partial and complete identity yields 94 languages with intensifier-reflexive identity, contrasting with 74 languages that do not exhibit identity in this domain (56 per cent versus 44 per cent). Although there is thus a slight preference for identity, this parameter divides our sample into two parts of almost equal size.

Language families manifesting mainly intensifier-reflexive identity are Sino-Tibetan, Dravidian, Afro-Asiatic, Altaic, Trans-New Guinea, and the Indic branch of the Indo-European family. Differentiation predominates in Uto-Aztecan, Penutian, and in the European branches of Indo-European. Among the language families that have members illustrating both identity and differentiation we find Austronesian, Australian, and Niger-Congo.

In terms of areal distribution there is a major divide, with Central Asia, East and South-East Asia, the South Asian subcontinent, and New Guinea showing a greater than average share of languages with identity, whereas Europe and the Americas mostly harbour languages that keep intensifiers and reflexives apart. Of course, the areal clustering in Europe can be attributed to genetic factors, but we seem to be dealing with a true *Sprachbund* phenomenon in Asia and in New Guinea. Identity is also found on the fringes of Europe, northern Africa, and the Middle East. Mesoamerica, central Africa, and Australia show no clear areal patterning.

4 Lexical sources

Another parameter that we have tried to investigate concerns the lexical sources of intensifiers and reflexives. However, due to a multitude of adverse factors (the lexical source is frequently not

transparent, etymological dictionaries exist for only a few languages, grammar books only rarely give this kind of information) we have only been able to find this information for slightly over sixty languages. Therefore no separate map is given. The languages for which we do have reliable information show that intensifiers and reflexives typically derive from expressions for body parts (also including 'soul'). Whether or not intensifiers are invariably an intermediate step in the development of reflexives from expressions for body parts is not clear in all cases.

Of the 62 languages for which we have information about the lexical source of the reflexive, 47 have expressions that are related to a body-part noun:

body:	30 languages (e.g. Igbo, Japanese, Evenki)
head:	12 languages (e.g. Abkhaz, Podoko)
soul:	2 languages (e.g. Modern Standard Arabic)
bone:	1 language (Modern Hebrew)
heart:	1 language (Dongolese Nubian)
skin:	1 language (Ngiti)

A high proportion of these are African languages. Other lexical sources include anticomitative expressions like 'alone' (Lealao Chinantec), expressions indicating an extreme or a precise value 'very, exact' (Chalcatongo Mixtec), 'return' (Paamese), and 'reflection in water' (Finnish).

Of the 62 languages with a known lexical source, 33 have reflexive-intensifier identity, and in 24 of these 33 languages the intensifier/reflexive is related to a body-part expression. Examples of languages whose reflexive pronoun, though distinct from the intensifier, is related to a body-part expression include Basque (*buru-* 'head') and Ngiti (*ndɨ* 'skin'). Conversely, there are languages that have an intensifier related to a body-part expression, but whose reflexive is derived from a different source (Rama: intensifier *páin* 'body', reflexive *áp* 'alone'). There are two closely related languages in the sample that draw a formal distinction between reflexives and intensifiers, such that both of them derive from expressions for body parts, albeit different ones: Koyraboro Senni intensifier *hundey* 'soul' versus reflexive *boŋ* 'head', Koyra Chiini intensifier *huneyno* 'body' versus reflexive *bomo* 'head'.

5 Implicational connections

Distinguishing two types of languages on the basis of the relationship between intensifiers and reflexives is interesting and relevant in that identity versus differentiation of intensifier and reflexive can be shown to correlate with other properties of reflexive pronouns. If a language uses the same expression both as intensifier and as reflexive pronoun, this expression is not used as a middle marker or marker of derived intransitivity (cf. König and Siemund 2000a: 59). Languages belonging to the same type as English never use reflexive markers in constructions like (7a–c) from German or (8a–b) from Spanish:

(7) German

a. *Karl* *setzte* *sich*.
Charles sat REFL
'Charles sat down.'

b. *Die* *Tür* *öffnete* *sich*.
the door opened REFL
'The door opened.'

c. *Dieses* *Hemd* *wäscht* *sich* *gut*.
this shirt washes REFL well
'This shirt washes well.'

(8) Spanish

a. *Se* *venden* *coches* *usados*.
REFL sell cars used
'Used cars for sale.'

b. *Aquí* *se* *habla* *español*.
here REFL speak Spanish
'Spanish spoken here.'

Other implicational connections that have been proposed and discussed in the literature relate to the possible antecedents of reflexives and to their binding domain. Thus reflexives that allow long-distance binding seem invariably to be different from intensifiers (Norwegian *seg*; Latin *se*, *sibi*), whereas reflexives of the English type can only be bound in a local domain. On the other hand, expressions that can only be used as reflexives seem to require a subject as antecedent, while expressions like English x-*self* also allow objects as antecedents. Such implicational connections have not been shown to hold, however, for a sufficiently large sample of languages.

6 The development of reflexive markers

Body-part expressions can develop directly into reflexive pronouns as a result of metonymic change (cf. Faltz 1985). Intensifiers also play an important role in the genesis and renewal of reflexive markers. In example (9) from Old English (Norman 1848), the addition of the intensifier to the referentially independent object pronoun makes this pronoun coreferential with the subject, leading to the development of a complete paradigm of reflexives (similarly to Malayalam).

(9) Old English
Judas *ahēng* *hine* *selfne*.
Judas hanged him self.ACC
'Judas hanged himself.'

Another way of renewing reflexive markers is by intensifying weak reflexives. Such developments are widely attested in Romance languages, where the reflexive clitics are being replaced by the respective tonic forms in combination with intensifiers (cf. (10) from Portuguese; similarly in Swedish, Norwegian, Dutch).

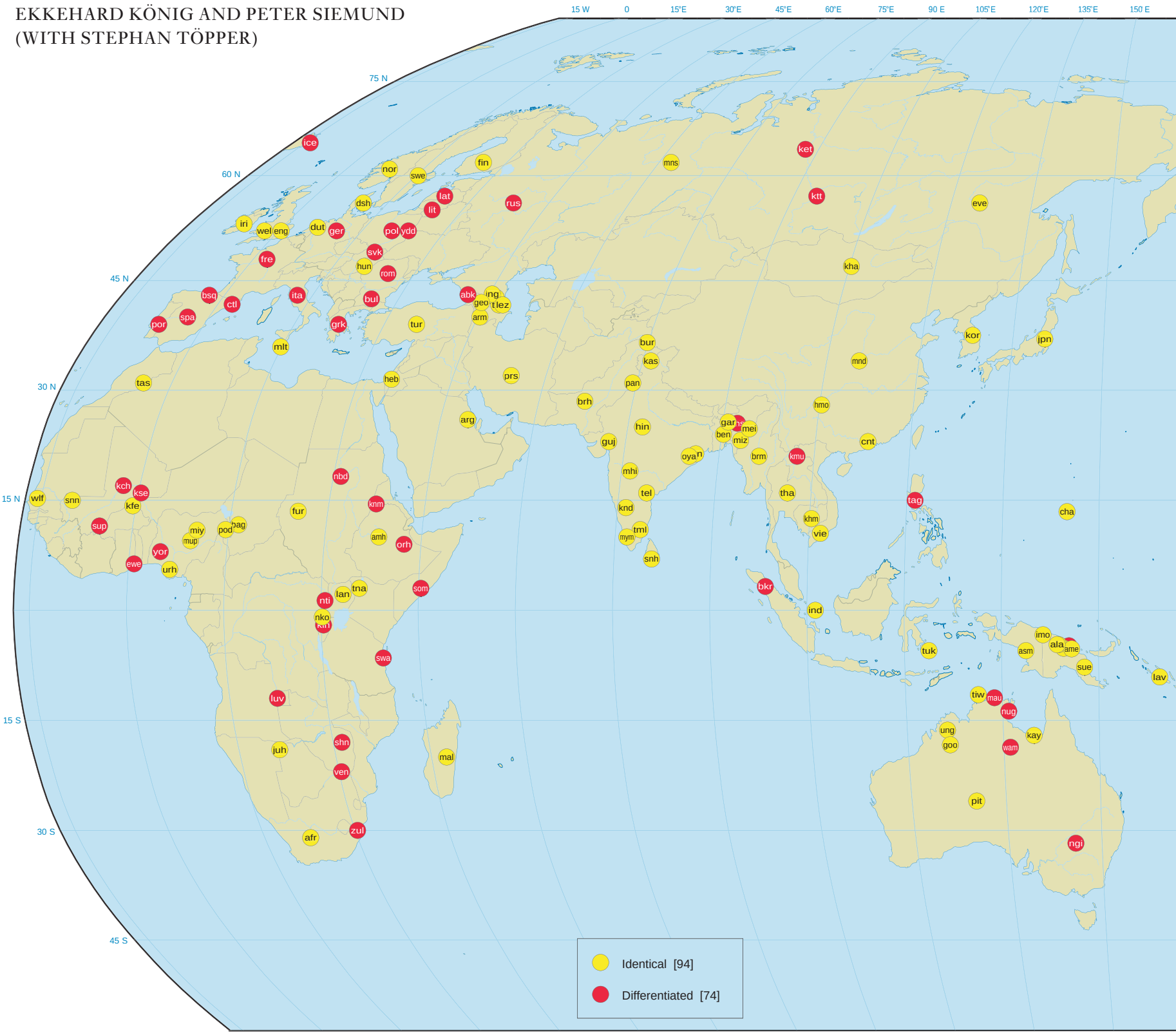
(10) Portuguese
Maria *olha* *a* *si* *mesma* *no espelho*.
Maria looks at REFL INTF in mirror
'Maria looks at herself in the mirror.'

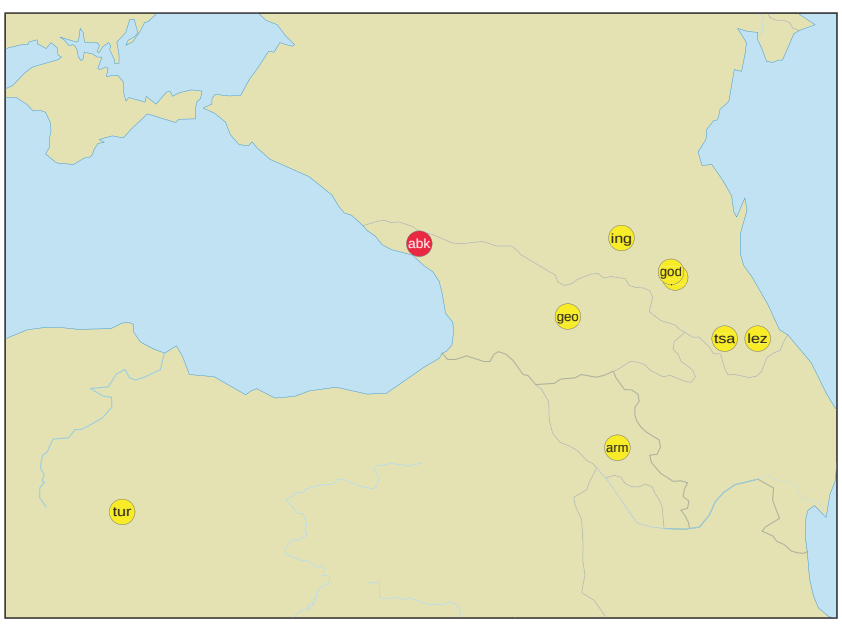
Finally, it is also possible that intensifiers by themselves come to be used as reflexives, as in Mezquital Otomí (Oto-Manguean; Mexico):

(11) Mezquital Otomí (Priego Montfort 1989: 120)
Bi *hyĕní* *sĕhĕ*.
3.PST cut INTF/REFL
'He cut himself.'

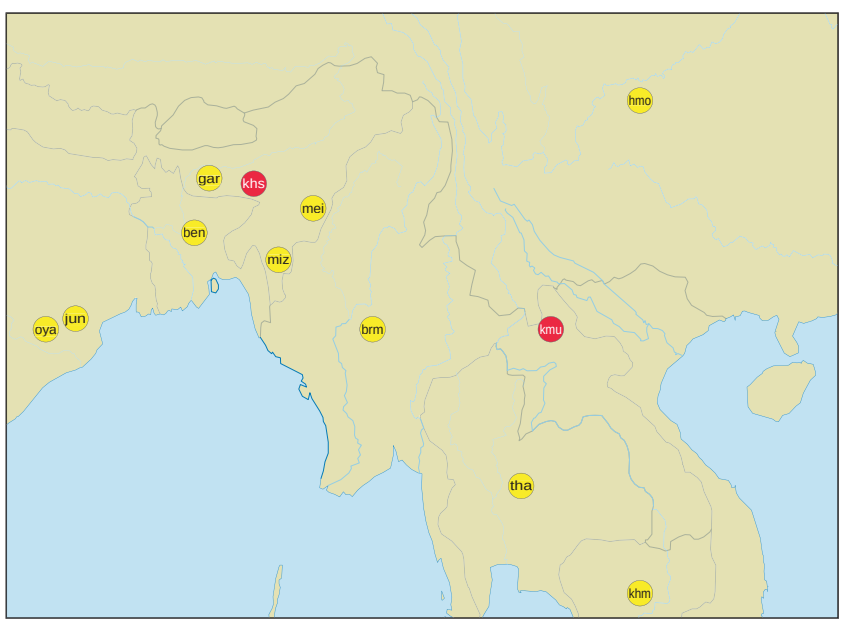
47 Intensifiers and Reflexive Pronouns

EKKEHARD KÖNIG AND PETER SIEMUND
(WITH STEPHAN TÖPPER)

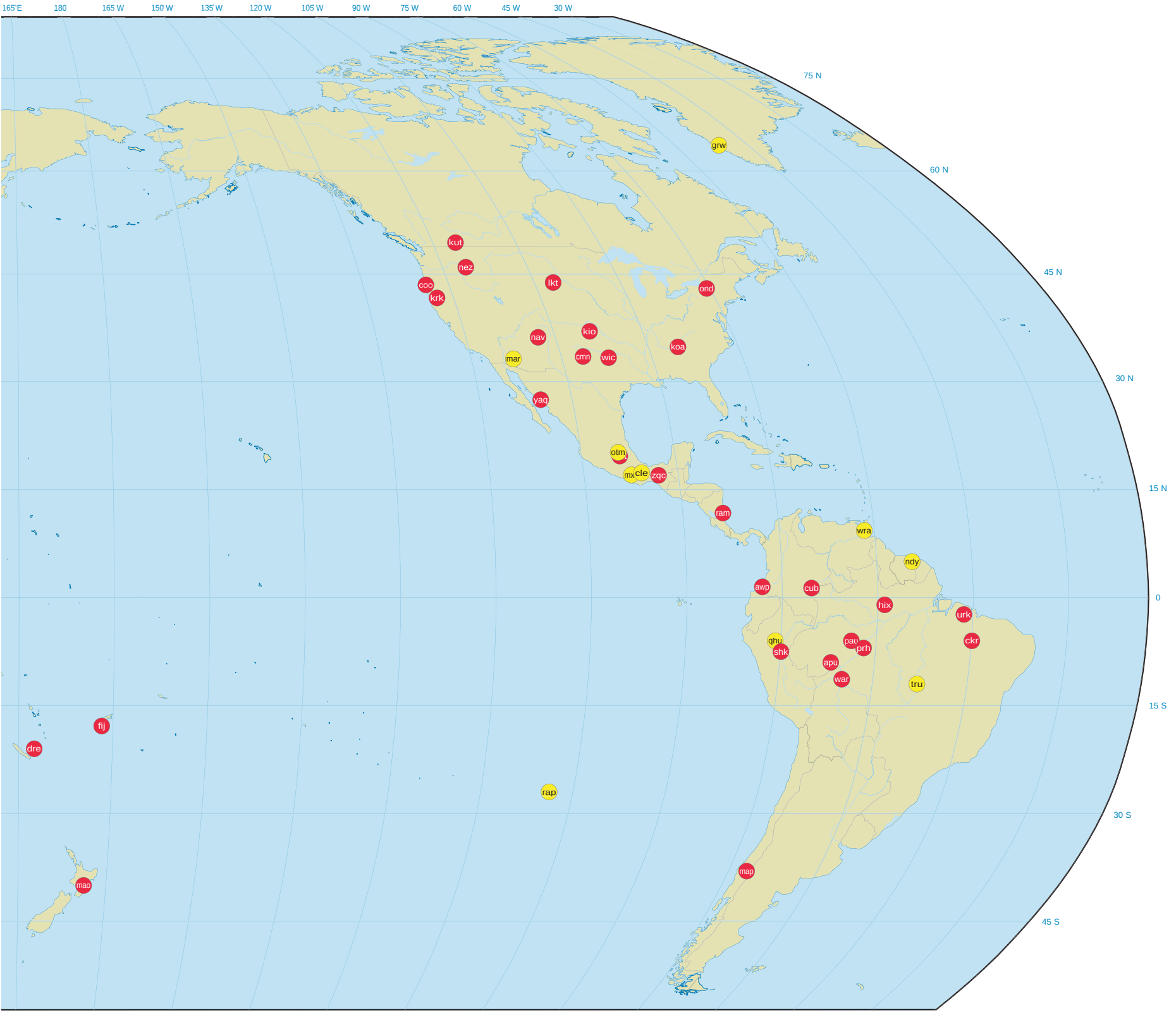


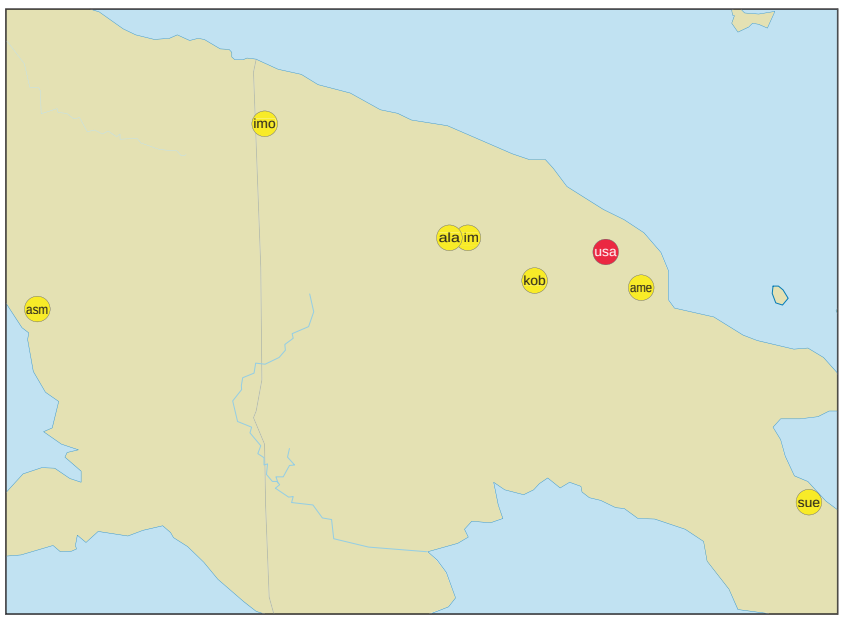


Caucasus and adjacent areas

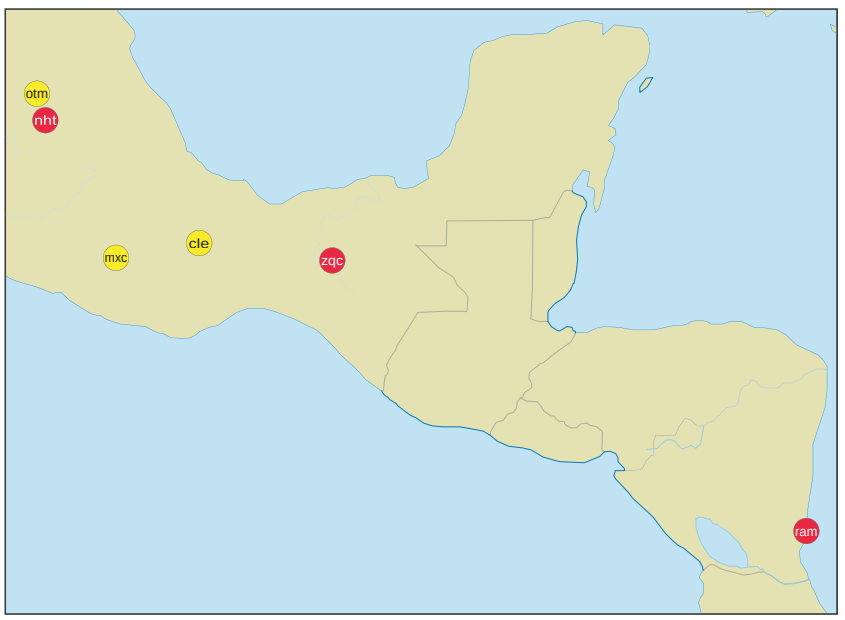


Myanmar and surrounding areas





central New Guinea



southern Mexico and Central America

48 Person Marking on Adpositions

DIK BAKKER

1 Defining the values

This map gives a survey of the distribution of languages with person markers on their adpositions. An example of a person-marked preposition is found in (1) from Maybrat (West Papuan; Papua, Indonesia).

(1) Maybrat (Dol 1999: 88)
T-ai *m-kah* *ara.*
1SG-hit 3SG.N-with stick
'I hit (it) with a stick.'

The 378 languages in the map have been characterized in terms of the four values in the table below.

1. No adpositions	63
2. Adpositions without person marking	209
3. Person marking for pronouns only	83
4. Person marking for pronouns and nouns	23
total	378

In order to determine whether a language has person marking on its adpositions, we first need to establish whether it possesses the category **adposition** at all. In general, the major function of an adposition is to relate its object, i.e. the noun phrase with which it forms a constituent, to another nominal or a verbal constituent on the basis of a more or less specific semantic relationship, such as location, time, property, instrument, or possession. Languages may use several strategies to fulfil this function, either lexical or morphological. In example (2) below, from Barbareño Chumash (California), we have a verbal construction fulfilling such a function. What is expressed by an adverb and an adposition in the English translation, *across from*, is rendered by a fully inflected verb in this language.

(2) Barbareño Chumash (Wash 2001: 75)
K^{h}-ili-ʔetemé.su"s *hi* *lwí.sa"* *hiklé-ḱen*
1-HAB-be.across.from DEP Luisa DEP-1.sit
hi-ho-·lamé.sa.
DEP-DIST-table
'I used to sit across from Luisa at the table.'

Example (3), from Tauya (Madang; Papua New Guinea), shows a nominal strategy to express the locality relationship rendered by the preposition *beside* in the English version.

(3) Tauya (MacDonald 1990: 283)
ya *nai-sa*
1SG rib-LOC
'beside me'

Although diachronically forms such as the verb *K^{h}iliʔetemé.su"s* in (2) and the noun *naisa* in (3) may eventually give rise to adpositions, they will not be considered as such here. In general a specific element in a language will be assumed to be an adposition only if it is morphologically independent and displays morphosyntactic behaviour distinct from more clear-cut verbal, nominal, or adverbial elements in that language.

In addition to languages such as Chumash and Tauya which encode adpositional meanings via verbs or nouns, there are languages in which the relevant relations are coded by morphological means, typically by a case suffix. The example from Arabana (Pama-Nyungan; South Australia) in (4) illustrates this strategy. The ablative case is used to express 'away from'.

(4) Arabana (Hercus 1994: 71)
Maka-ru *kilta-rnda.*
fire-ABL pull-PRES
'He pulls it out of the fire.'

Another type of morphologically dependent expression of nominal relations is via adverbial nominal affixes, as in example (5) from Jaqaru (Aymaran; Peru).

(5) Jaqaru (Hardman 2000: 21)
ut-nuri-t"a
house-within-from
'from within the house'

Arguably a borderline case between morphological and syntactic expression are clitics which attach to the noun phrase rather than to nouns, as exemplified in (6) from Ngankikurungkurr (Daly; Northern Territory, Australia).

(6) Ngankikurungkurr (Hoddinott and Kofod 1988: 72)
Kalla *ngayi* *yedi* *tye* *yeningkisyi* *yaga=nide.*
mother 1SG went PST canoe DEM=LOC
'My mother came in that canoe.'

In many instances, case affixes as in (4) and clitics as in (6) can be analysed diachronically as the result of the affixation or cliticization of adpositions. However, taking a rather conservative, strictly syntactic position on this matter, I will not consider such bound forms as true adpositions. A further argument for not doing so may be that since such forms are (syntactic) dependents rather than heads, they are not potential targets for person marking but, to the contrary, may attach themselves to pronominal forms. As a result of this, all languages which resort exclusively to strategies as exemplified under (2)–(6) are coded as having no adpositions (value one) on the map. They are by definition irrelevant for the type of person marking under consideration here.

Having established the criteria for what constitutes an adposition, we must now determine what qualifies as **person marking on adpositions**. In order for a potential marker to be considered a person marker it should fulfil two requirements. First, there should be enough differentiation between the relevant forms that a distinction exists either between all three persons or between any combination of first, second, or third person. One of the forms may be zero. And secondly, the forms should be affixes rather than clitics.

Languages with only bare adpositions, and those for which markers on their adpositions do not meet the criteria mentioned above irrespective of the nature of the adpositional object, are assigned value two (**adpositions without person marking**). English is a case in point, and so is Polish, as shown in example (7a–b).

(7) Polish (Anna Siewierska p.c.)
a. *Idę* *do* *Kasi.*
go.FUT.1SG to Kasia.GEN
'I go to Kasia.'
b. *Idę* *do* *niej.*
go.FUT.1SG to 3SG.F.GEN
'I go to her.'

When languages do show person marking on adpositions for pronominal objects but not for nominal objects, they are assigned value three (**person marking for pronouns only**). Paamese (Oceanic; Vanuatu) has this kind of adpositional marking, as shown in (8a–b) below.

(8) Paamese (Crowley 1982: 182)

a. *Mail Ham sān lēta min-nau*
Mail Ham 3SG.send letter to-1SG
ranaut Vīla.
from.place Vila
'Mail Ham sent me a letter from Vila.'

b. *Kai selūs min tāta ven mane onak.*
3SG 3SG.speak to father about money POSS.1SG
'He spoke to father about my money.'

Typically, in these languages the adposition has no separate independent pronominal object apart from the person marker itself, although there may be one under certain conditions, such as contrastive stress. Hence, more often than not, the person marker on the adposition is the only element representing the pronominal object. In this group are also included languages for which only a subset of the adpositions have person marking. In Burushaski (isolate; Pakistan or India) the postposition *paču* 'with' in (9a) occurs with a prefixed person marker; the postposition *gʌnɛ* 'for' in (9b) does not.

(9) Burushaski (Lorimer 1935: 96, 97)

a. *Ja a'-pʌču huru't.*
I 1SG-with stay
'Stay with me.'

b. *ja gʌnɛ.*
I.GEN for
'for my sake'

The occurrence of person marking may also be subject to other constraints. For instance, in Kiribati (Oceanic) there is person marking only when the referent is animal, as shown in (10a–b).

(10) Kiribati (Groves et al. 1985: 65)

a. *nako-ia mooa*
to-3PL chickens
'to the chickens'

b. *nakon taian nii*
to ART coconut.trees
'to the coconut trees'

However, borderline cases such as Spanish, which has bound pronominal forms for just one preposition, *con* 'with', and only for the first- and second-person singular forms, were coded as showing value two.

Finally, there are languages which have **adpositional marking with both pronominal and nominal objects**. Abkhaz (North-West Caucasian; Georgia) is a case in point, as shown in (11).

(11) Abkhaz (Hewitt 1979: 103)

a. *a-jəyas a-q'nə*
DEF-river 3SG-at
'at the river'

b. *sarà s-q'əntº'*
I 1SG-from
'from me'

Such languages are assigned value four. Like value three, this value includes languages which do not normally use a full pronoun in constructions such as (11b).

In my sample there were no languages of the other logically possible type, which would have person marking on adpositions for nominal but not for pronominal objects.

2 Geographical distribution

The phenomenon of person marking on adpositions is not distributed evenly over the world. First, the languages of North America and Australia lack adpositions much more often than the other areas. Of the North American languages in the sample, 48 per cent are adpositionless; for Australia the percentage is as high as 73 per cent. The overall figure for adpositionlessness is around 17 per cent in the current sample.

If we look only at languages with adpositions, then the following may be observed. Person marking on adpositions is nonexistent in Australia, and rare in South-East Asia (three out of thirty languages, all belonging to the Austronesian family). For South-East Asia this does not come as a surprise, given the general lack of person marking in this area. In Australia, on the other hand, person agreement on the verb is found in the majority of the languages, which may be an indication that the two types of person marking are not necessarily closely related to each other. In Mesoamerica and the Pacific, person marking on adpositions is clearly abundant. These are the only areas where a majority of the languages have this kind of marking: over 60 per cent.

Other striking figures are the relatively frequent occurrence of languages of type four (person marking for pronouns and nouns) in Mesoamerica. It is above all the Mayan and Uto-Aztecan languages which have this property. In Africa, although this area has a relatively high amount of type three languages, none of the languages show type four. Eurasia, New Guinea, and South America are all close to the overall distribution. Of the twenty Indo-European languages in the sample, only the two Celtic languages Irish and Welsh and the two Iranian languages Persian and Kurdish have any marking at all; they all show type three.

3 Theoretical issues

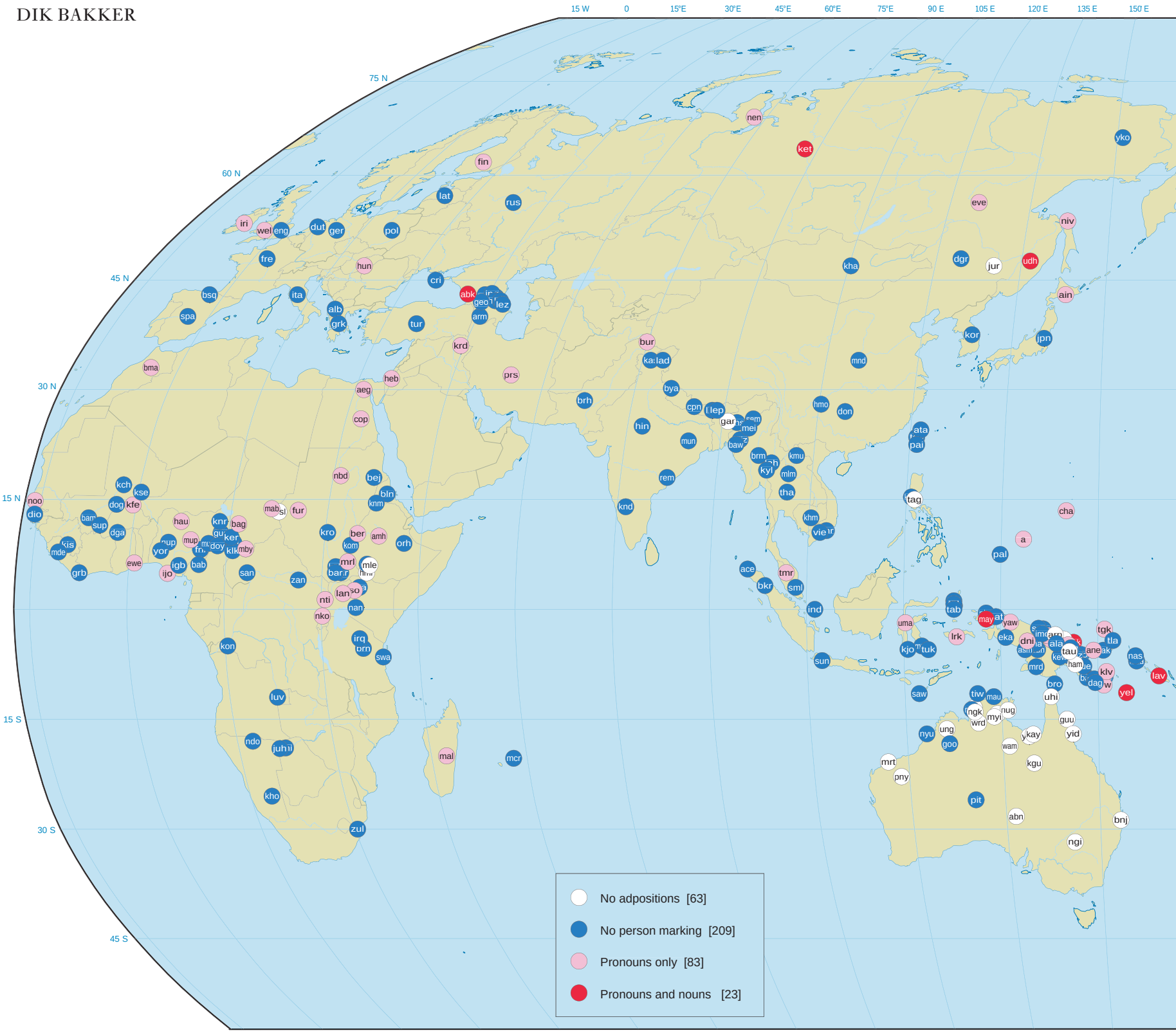
Due to the fact that linguists differ in the criteria that they use in determining the existence of adpositions in a language, there may be considerable discrepancies in the percentage of languages which are seen to lack adpositions altogether. My figure of 17 per cent is close to that of Hawkins (1983), and also Tsunoda et al. (1995). By contrast, Matthew Dryer (Chapter 85 of this atlas) is more liberal in his interpretation of what constitutes an adposition and treats in some cases as adpositions what I have considered to be case clitics.

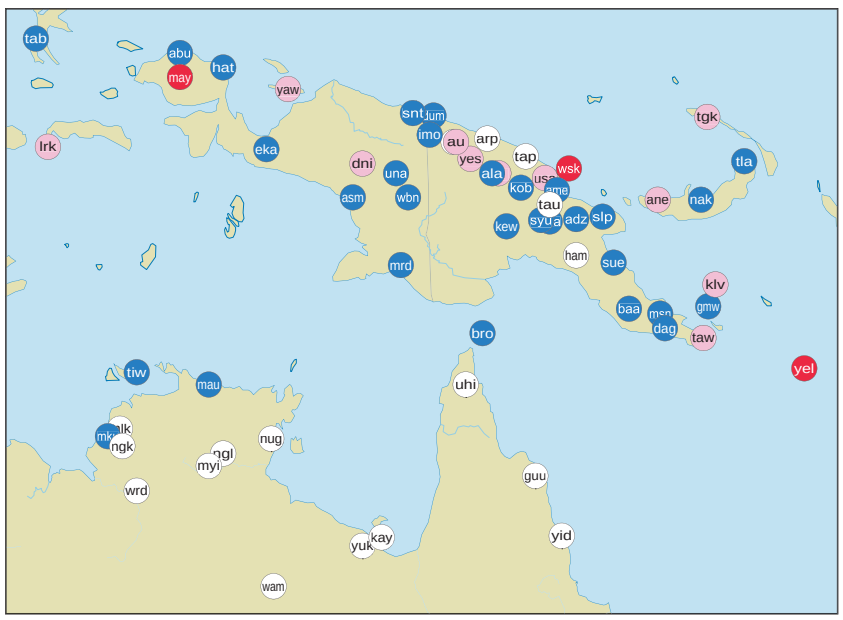
A second point that deserves a word of comment is that it is not always possible on the basis of the information presented in grammars to decide what the precise grammatical status of a specific person marker is, i.e. whether it is indeed an affix attached to the adposition, or a clitic, or even a more or less independent pronominal form. Apart from morphosyntactic information, the phonological form of the person marker may be an indication, as well as the amount of (dis)similarity to other person markers, notably independent and possessive pronouns.

A final issue which needs to be briefly mentioned is the diachronic status of the markers. They may be relatively recent, and introduced after adpositions were formed in the language. In these cases they are often phonologically related to the actual personal pronouns of the language, typically in their object or oblique form. This is the case in about 12 per cent of the languages that display some form of person marking on their adpositions. Or the markers may be relatively old, and already present on the verbal or nominal precursor of the adposition. If there is similarity to any other person marker in the language at all, similarity to subject or object agreement markers on the verb may (but need not) be an indication of the verbal origin of the adposition. This seems to be the case in about 13 per cent of the relevant languages. Similarity to possessive markers, which was attested in about 24 per cent of the relevant languages, may point to a nominal origin. In around 41 per cent of the cases, person markers on adpositions bear a similarity to both verb agreement and possessive markers in the language concerned. In the remaining 11 per cent no clear similarities could be detected.

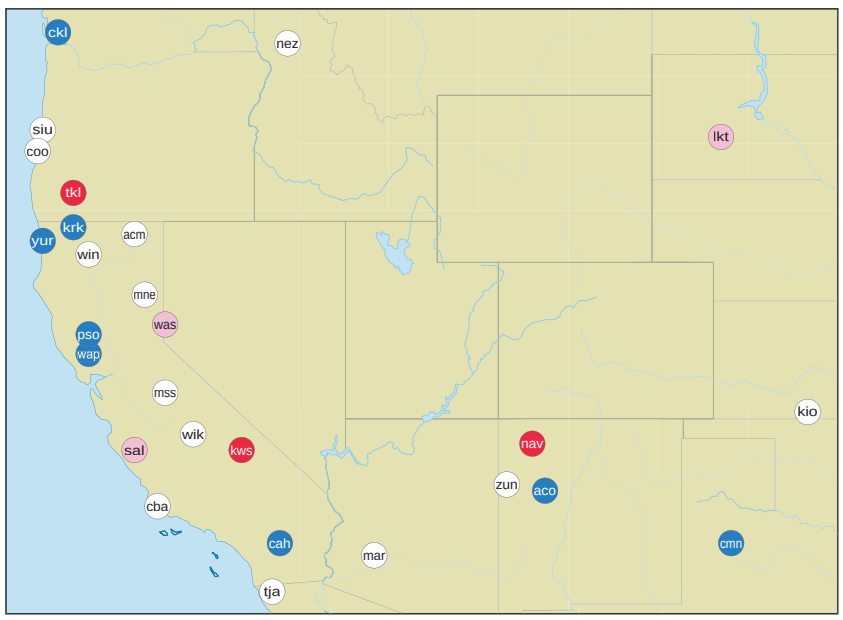
48 Person Marking on Adpositions

DIK BAKKER

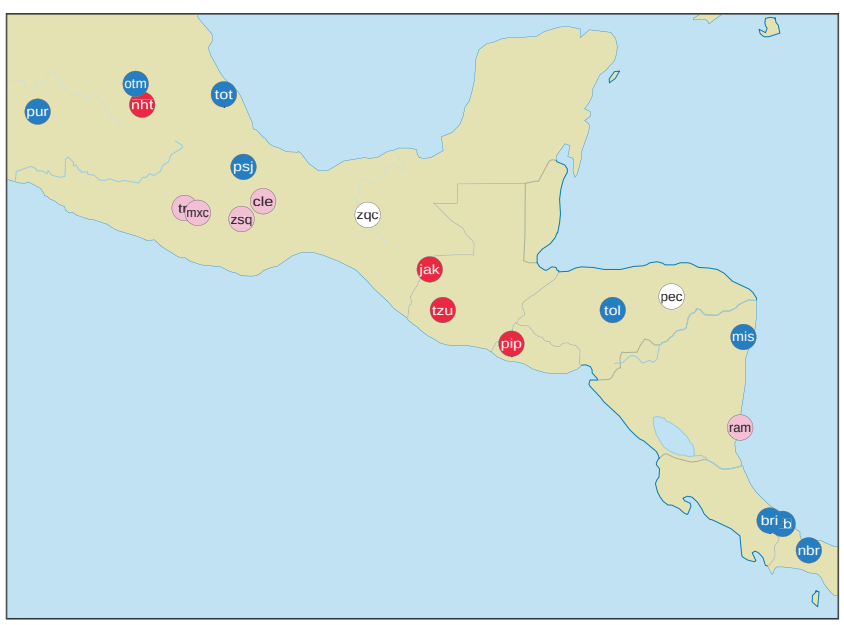




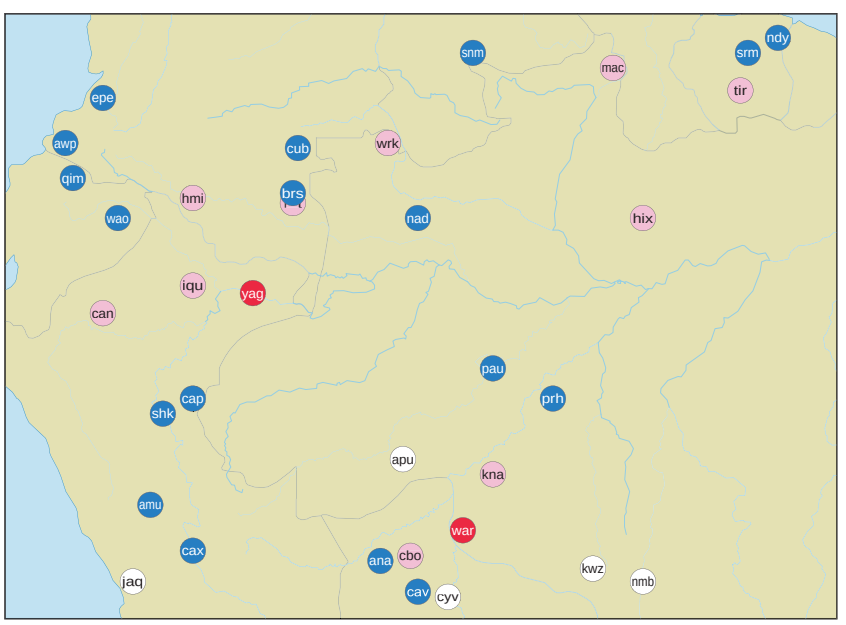
New Guinea and northern Australia



western United States



southern Mexico and Central America



northwestern South America

49 Number of Cases

OLIVER A. IGGESEN

1 Defining the values

Morphological case on nominals is a common device to express the syntactic and semantic relationships between clausal constituents. However, the languages of the world that use this strategy vary greatly with respect to the number of case categories represented in their inflectional system. It is the purpose of this map to display this numerical variation in the **productive** case paradigms of **substantives** (full nouns) in 261 languages. Deviant case inventories found in minor word classes of particular languages are dealt with in Chapter 50.

○	1. No morphological case marking	100
○	2. 2 case categories	23
○	3. 3 case categories	9
○	4. 4 case categories	9
○	5. 5 case categories	12
○	6. 6–7 case categories	37
○	7. 8–9 case categories	23
○	8. 10 or more case categories	24
◇	9. Exclusively borderline morphological case marking	24
	total	261

The feature values are largely self-explanatory. In the languages lacking morphological case (e.g. Vietnamese), grammatical relations are expressed by word order and/or morphologically and prosodically independent function words (in general, prepositions and postpositions), and partly also by morphological devices on the verb.

The minimal case paradigm contains **two members**, since **paradigmatic relationships** between word forms are ultimately based on **binary oppositions** (minimal pairs). This implies that whenever a language has an overtly marked case category expressing a specific function, a corresponding zero-marked base form is counted as a case ("default case" or "direct case") even if it has no specific function describable in positive terms. In such instances, the base form receives its case status only by virtue of contrasting with a functionally and formally marked case category. An example in point is Mapudungun (Araucanian; Chile), which has only one overt case suffix *-mew* ~ *-mu* expressing diverse oblique functions such as place, cause, and instrument. This suffix also occurs on syntactically demoted core arguments. Hence, Mapudungun exhibits the following case system, exemplified by the word *peñi* 'brother'.

(1) Mapudungun (Smeets 1989: 77)

Direct:	*peñi*
Oblique:	*peñi-mu*

A language with **three cases** is the Obdorsk dialect of Khanty (Uralic; western Siberia). Apart from the unmarked direct case, there is a locative and a translative (encoding the goal of movement and change of state as well as secondary predicates). The model word is *xo:t* 'house'.

(2) Obdorsk Khanty (Nikolaeva 1999: 13)

Direct:	*xo:t*
Locative:	*xo:t-na*
Translative:	*xo:t-ti*

While two or three cases constitute small paradigms, languages with four or five categories may be considered to have mid-sized inventories. **Four cases** are found in Icelandic. The table lists the paradigm of *hestur* 'horse'.

(3) Icelandic (Thráinsson 1994: 153)

Nominative:	*hest-ur*
Accusative:	*hest*
Genitive:	*hest-s*
Dative:	*hest-i*

An example of a **five-case** language is Trumai (isolate; Mato Grosso, Brazil). Trumai has two different dative markers applying to full nouns: *-atl*, used with identifiable referents, and *-ki*, used with less identifiable referents. Although the selection of the two endings is clearly governed by semantic criteria, this meaning difference does not affect the case relation per se expressed by the markers; hence, there is only one dative category in Trumai. The example word is *axos* 'child' except in the locative, which does not occur with animate nouns in this language, and is thus represented here by *esak* 'hammock'.

(4) Trumai (Raquel Guirardello, p. c.)

Absolutive:	*axos*
Ergative:	*axos-ak*
Dative:	*axos-atl, axos-ki*
Genitive:	*axos-kate*
Locative:	*(esak-en)*

Languages with six through nine case categories have large inventories. Russian exemplifies the value **six to seven cases**. It has six productive cases (there are secondary genitives and locatives occurring only in some declension classes; these have not been judged productive). Russian case inflection is characterized by a high degree of declension-specific allomorphy and syncretism (see Chapter 28 for case syncretism). Two word paradigms belonging to two different declensions (*zavod* 'factory' and *karta* 'map') are given here.

(5) Russian (Wade 1992: 53, 69)

Nominative:	*zavod*	*kart-a*
Accusative:	*zavod*	*kart-u*
Genitive:	*zavod-a*	*kart-y*
Dative:	*zavod-u*	*kart-e*
Instrumental:	*zavod-om*	*kart-oj*
Locative:	*zavod-e*	*kart-e*

The feature value of **eight to nine cases** is represented by West Greenlandic (Eskimo-Aleut). The model word is *qimmiq* 'dog'.

(6) West Greenlandic (Fortescue 1984: 206)

Absolutive:	*qimmiq*
Ergative:	*qimmi-p*
Instrumental:	*qimmi-mik*
Allative:	*qimmi-mut*
Locative:	*qimmi-mi*
Ablative:	*qimmi-mit*
Prosecutive:	*qimmi-kkut*
Equative:	*qimmi-tut*

Languages with **ten or more cases** show very large paradigms. The languages on the map with the largest paradigms are Hungarian with (under some analyses) 21 productive cases, followed by Kayardild (Tangkic; Queensland, Australia) with 20, and Lak (Nakh-Daghestanian; eastern Caucasus) with 19 cases. The table of example (7) presents the (reasonably) productive case categories of Hungarian on the basis of the example word *hajó* 'ship'.

(7) Hungarian (Tompa 1968: 206–9)

Nominative:	*hajó*
Accusative:	*hajó-t*
Inessive:	*hajó-ban*
Elative:	*hajó-ból*

Illative:	*hajó-ba*
Superessive:	*hajó-n*
Delative:	*hajó-ról*
Sublative:	*hajó-ra*
Adessive:	*hajó-nál*
Ablative:	*hajó-tól*
Allative:	*hajó-hoz*
Terminative:	*hajó-ig*
Dative:	*hajó-nak*
Instrumental-Comitative:	*hajó-val*
Formal:	*hajó-képp*
Essive:	*hajó-ul*
Essive-Formal(-Similitive):	*hajó-ként*
Translative-Factitive:	*hajó-vá*
Causal-Final:	*hajó-ért*
Distributive:	*hajó-nként*
Sociative:	*hajó-stul*

The feature value **exclusively borderline case marking** refers to languages which have overt marking only for concrete (or "peripheral" or "semantic") case relations, such as locatives or instrumentals. This type is represented by Plains Cree (Algonquian; Saskatchewan, Alberta), whose only case-inflecting device is the locative suffix *-ehk* (Wolfart 1973: 31). In several descriptive traditions such paradigmatically isolated adverbial categories are considered derivational rather than inflectional. The functional core of case morphology is the expression of the specific syntactic relations of clausal arguments. Hence, languages like Cree do not participate in "genuine" case marking. In practical terms, delimiting case markers from other morphological-functional entities is often particularly difficult in such languages. However, a rather generous stance is adopted for the present map: as soon as one single morphological category is significantly involved in some kind of argument encoding, the entire set of forms is considered a "genuine" case paradigm. Thus Abkhaz (North-West Caucasian; Georgia) is listed as a two-case language because its adverbial case suffix *-s* is used to render secondary predicates ('to consider somebody as . . .') and thereby has a grammatical function, even if its impact on the overall argument structure of the language is somewhat marginal (Hewitt 1989: 101). It must be pointed out that full case-marking paradigms may contain categories that would be considered borderline cases if taken in isolation. Thus the borderline feature value has rather to be viewed as a threshold beyond which all categories, regardless of their semantics, are counted as true cases.

2 Theoretical issues

For the purposes of this chapter Blake's (1994: 1) general definition of case has been adopted:

> **Case** is a system of marking dependent nouns for the type of relationship they bear to their heads. Traditionally the term refers to inflectional marking, and, typically, case marks the relationship of a noun to a verb at the clause level or of a noun to a preposition, postposition or another noun at the phrase level.

Since case is by definition **dependent marking**, **head-marking** nominal morphology remains outside the limits of case even if it encodes case-like functions (e.g. possessive marking on the possessed instead of the possessor noun; see Chapters 23–5).

Morphological case is a combination of **form** and **function**, but distinctiveness of formal expression will be taken as criterial here. Hence, if in a language two or more unrelated functions are consistently expressed with the same marker, this is counted as only one case. The complete **syncretism** of dative and locative in Serbian-Croatian is thus considered as a single category. On the other hand, **allomorphic** case markers and **free variants** are not counted as separate cases as long as they are not associated with a discernible difference in meaning. In allomorphy-rich languages like Russian (see (5)) the general case paradigm is found by comparing the distinctive distribution patterns of markers across all productive declension classes. If in a language animate and inanimate nouns show divergent case inventories (as in Trumai (4)), the cases are added together even though no nominal is capable of inflecting for all of them.

The functional part of the above definition entails a number of stipulations. Categories with pragmatic (nonsyntactic) functions, such as **vocatives** or **topic markers**, are not counted as case even if they are morphologically integrated into case paradigms. **Genitives** are counted as long as they do not encode categories of the possessum like number or gender as well, if they do not show explicit adjective-like properties. However, genitives may take additional case affixes agreeing with the head noun case ("double case", Plank 1995); in that case they are not regarded as adjectival. **Case stacking**, i.e. the obligatory affixation of certain case markers to already case-inflected bases (e.g. in Imonda (Border family; Papua New Guinea; Seiler 1985: 73), where the ablative marker is added to the locative form: *sagl-ia-nèi* 'from the festival'), has been dealt with by counting the markers separately if both markers contribute to the combined meaning of the resulting word form (cf. Comrie 1999). Case-like markers with **derivational** character are not taken into account. This excludes, for example, the "locative" suffixes in Oneida (Iroquoian; Ontario) from being counted as case, since they can derive body-part nouns which may occur in all semantically permitted syntactic positions (i.e. not only as locational adverbials).

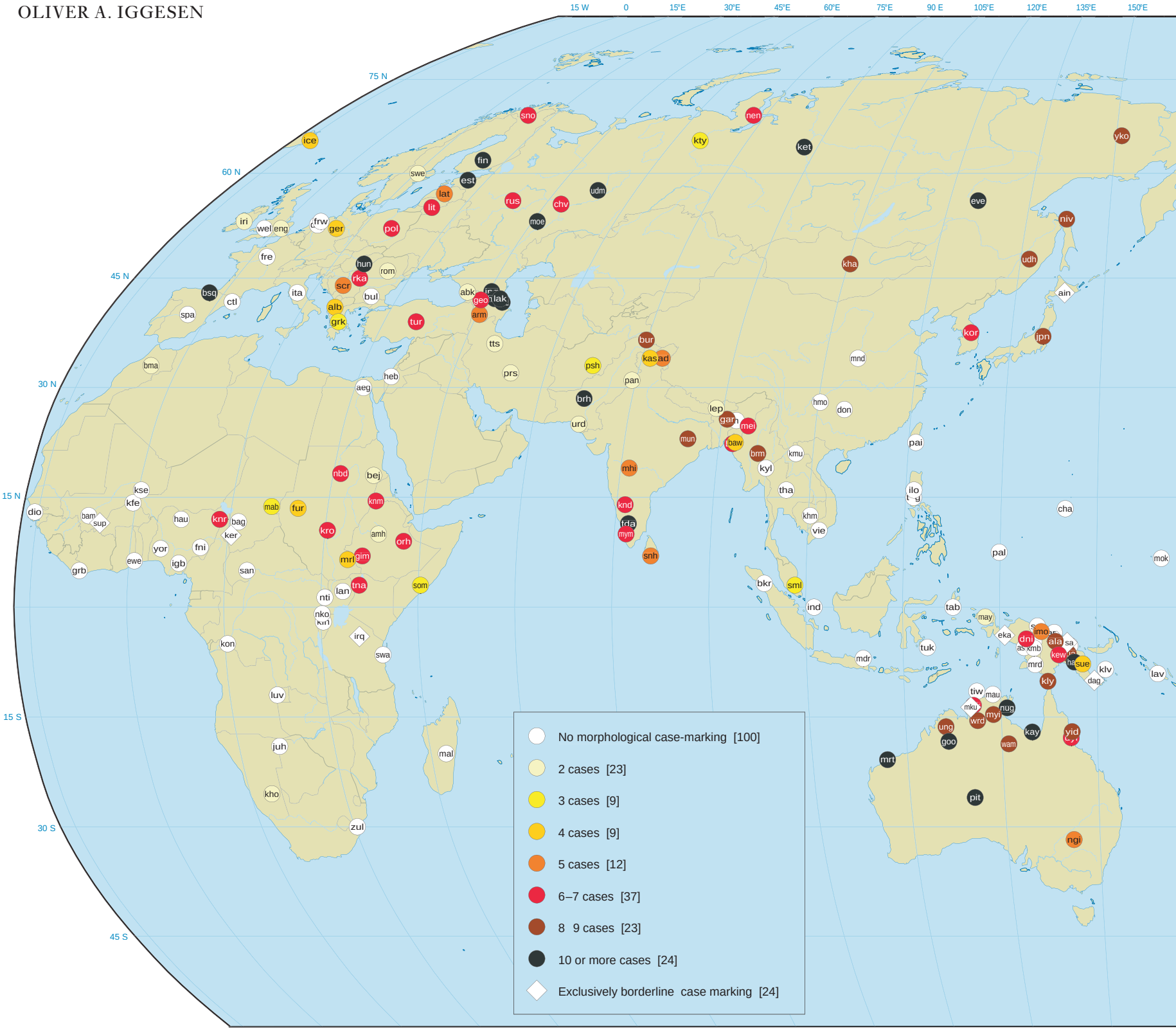
In a departure from Blake and partly also from Chapter 51 of this atlas, a more generous stance has been taken with regard to **clitic** and **adpositional** case marking. A marker does not necessarily have to be attached to the phrasal head to be counted as nominal case; it is only required that the marker show a sufficient degree of **bondedness** (phonological integration) with its host noun in **basic syntactic constructions**—i.e. in nonexpanded, head-only NPs. The reason for doing this is that postpositions (independent words), phrasal clitics, and inflectional case morphemes are diachronically interconnected on a grammaticalization cline, and it seems rather arbitrary to set up cut-off points on it. Furthermore, it is often problematic to decide whether in a given language a prespecified cut-off point has or has not been crossed (given the particular morphosyntactic properties of the language or the shortcomings of the extant descriptions). Finally, taking the semantic side of the case categories into account, there is little reason to keep clitic case marking separated from affixal marking only because of gradual differences in bondedness, while functionally categories in two languages representing different morphological types may be straightforwardly comparable.

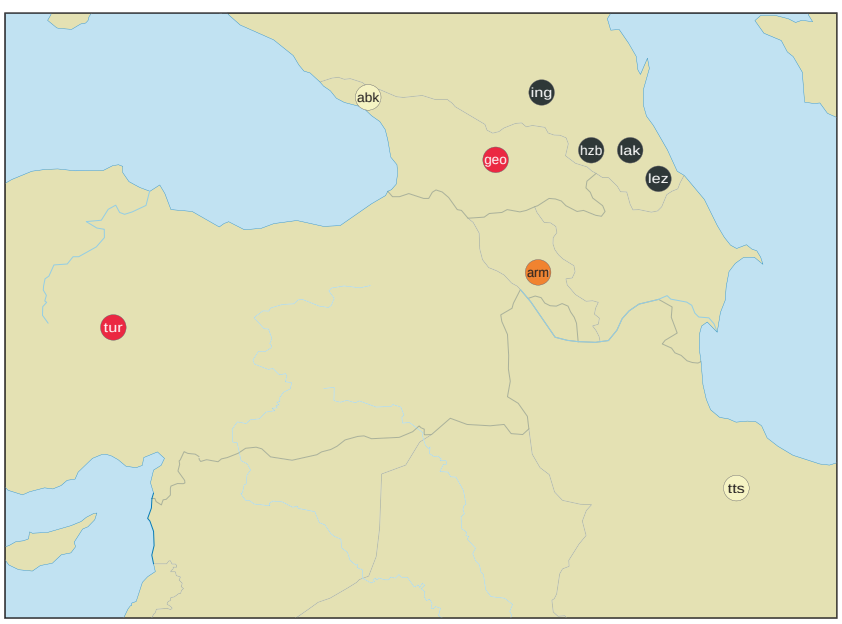
3 Geographical distribution

The map reveals some significant areal patterns. Large and very large inventories are common in northern and central Eurasia (extending to the Eskimo languages of North America), Australia, and the Caucasus. South Asia is characterized by mid-sized or large inventories; and large or mid-sized systems are also typical of eastern Europe, east-central Africa (i.e. Nilo-Saharan and Cushitic languages), and California. Borderline case-marking systems occur especially often in South and North America, New Guinea, and occasionally in the central belt of Africa. Western Europe has no case marking, or only minimal inventories (with the striking exception of Basque). Caselessness is even more typical of South-East Asia and the Pacific, and is also widespread in large parts of sub-Saharan Africa and in Mesoamerica. Perhaps the least uniform macro-area is South America (where large inventories show a regional concentration in the Andean and sub-Andean areas), but New Guinea too is quite diverse.

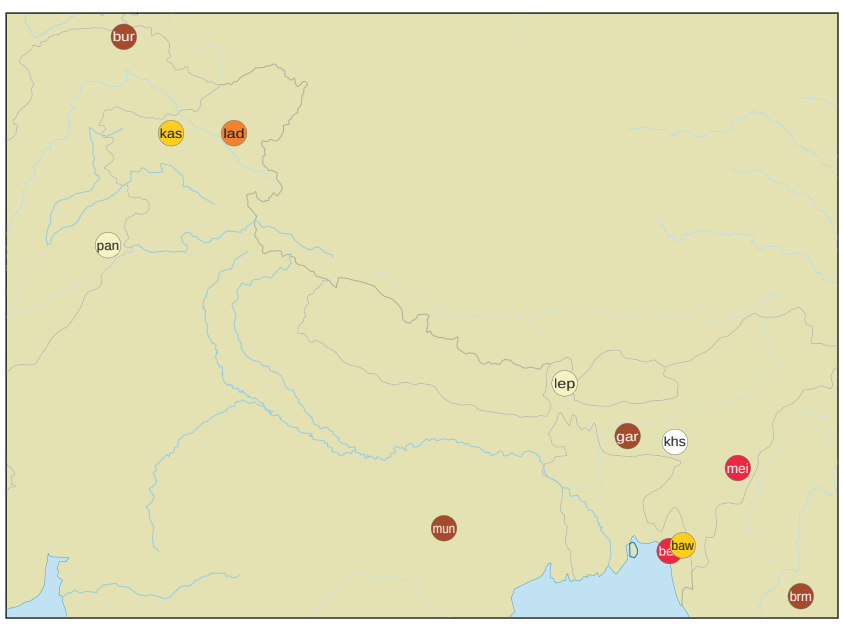
49 Number of Cases

OLIVER A. IGGESEN

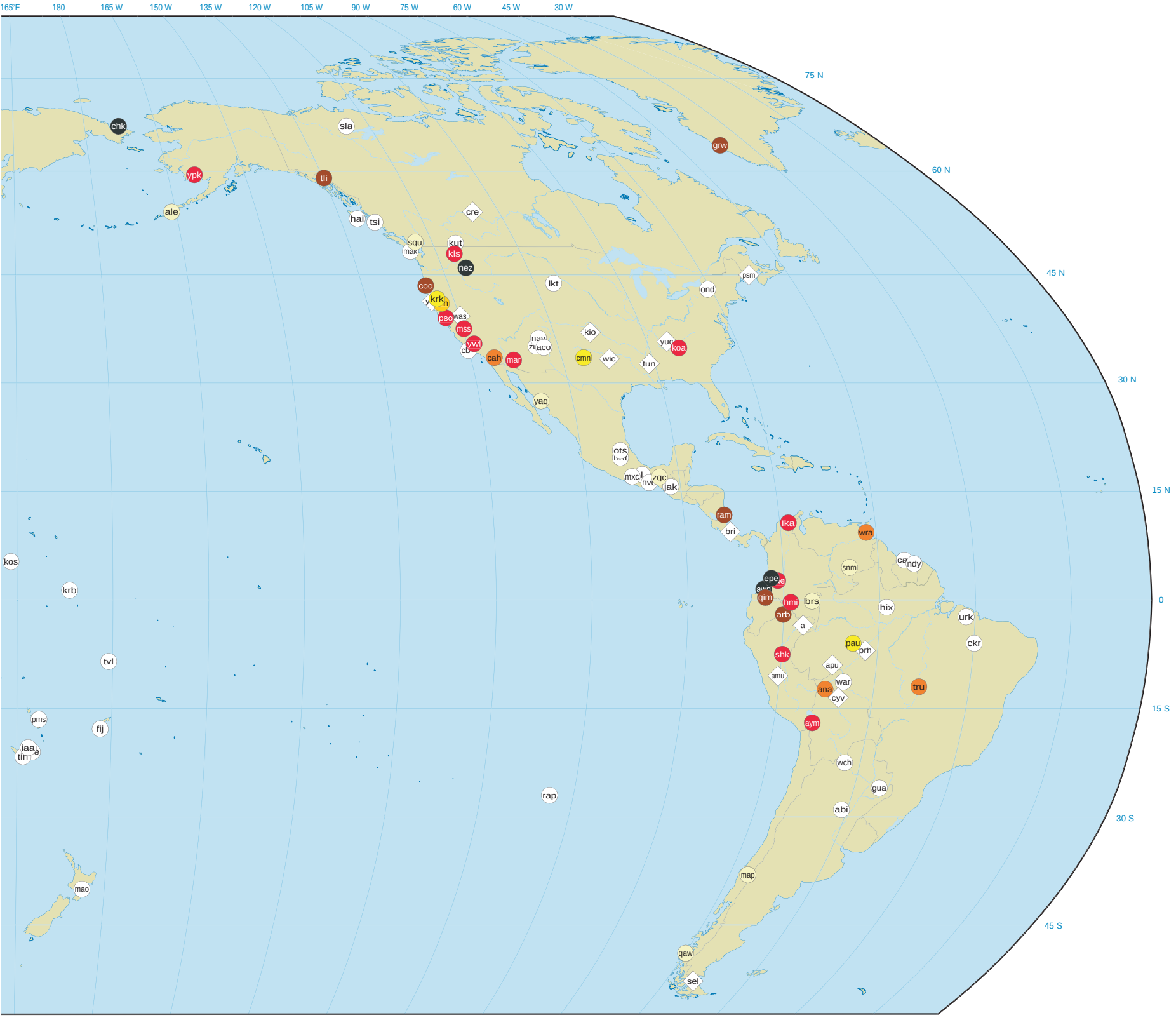


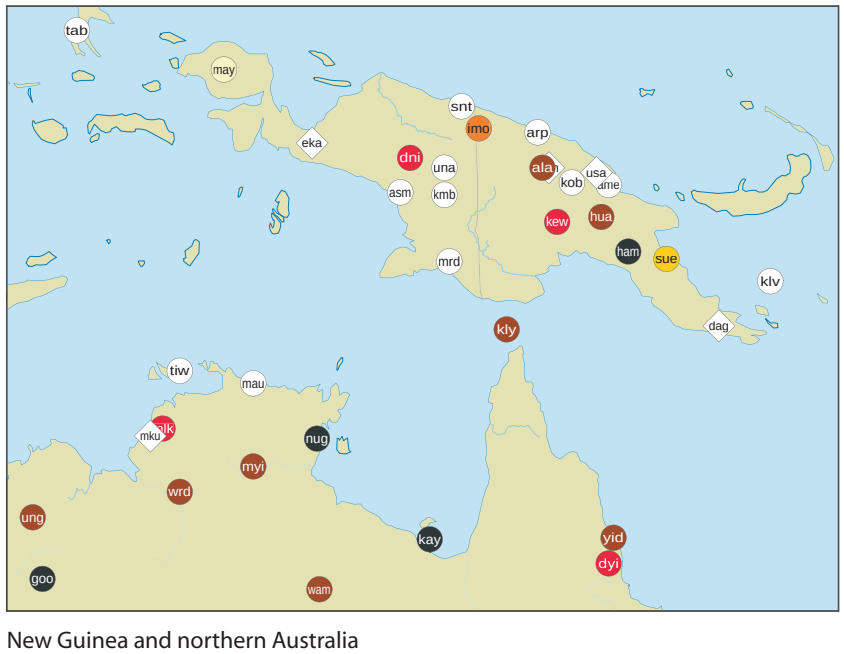


Caucasus and adjacent areas

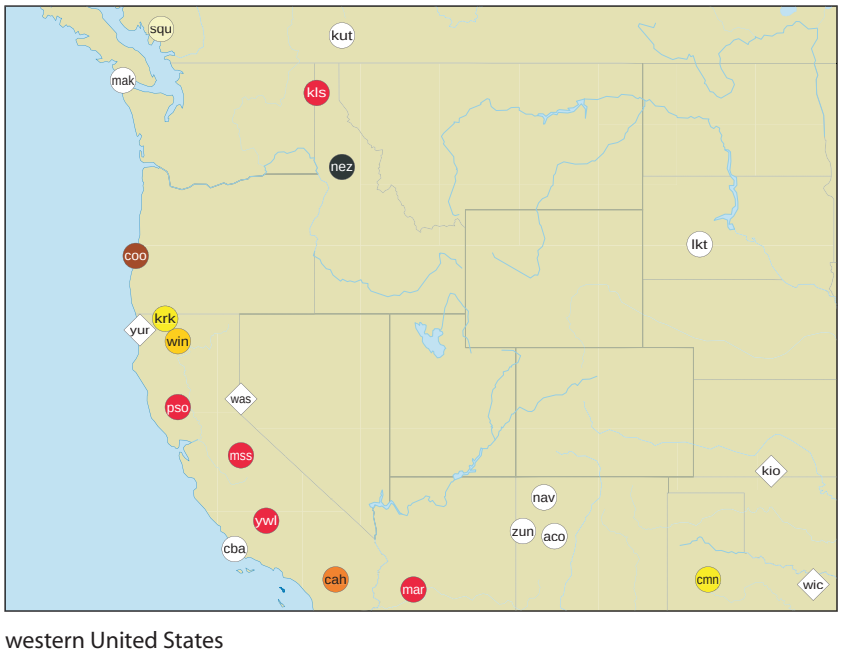


Himalayas and adjacent areas





New Guinea and northern Australia



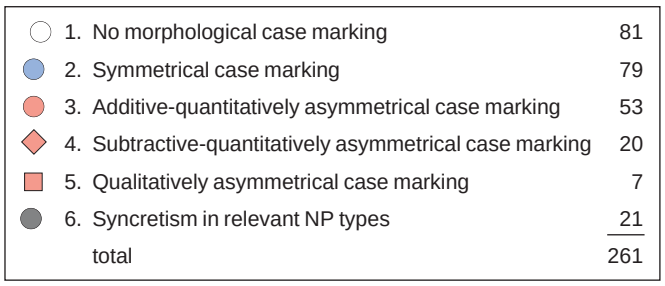
western United States

50 Asymmetrical Case Marking

OLIVER A. IGGESEN

1 Defining the values

Morphological case inflection is an efficient strategy for encoding syntactic relationships. Therefore one might naively expect that case-marking languages should use it consistently, applying the same case categories to all nominal expressions occurring in the appropriate syntactic positions. Indeed, there are many case-marking languages of this kind. Such languages exhibiting identical sets of case categories for all nominals are called **case symmetrical**. Other languages, however, apply case marking more selectively across their lexicon, restricting the occurrence of certain (or all) cases to a subset of their nominals. The different subsets of nominals therefore display different case inventories. This typological property of languages is called **case asymmetry**, and the languages themselves, **case asymmetrical**. This map shows the distribution of case asymmetry versus case symmetry on the basis of 261 languages (the same sample as in Chapter 49), with case-asymmetrical languages being further divided into distinct subtypes.



The phenomenon of case asymmetry is based on the division of a language's nominals into subclasses sharing common semantic or functional characteristics (i.e. into nonmorphologically based natural classes); these subclasses are called **NP types**. Case asymmetry manifests itself as differences between the case inventories of different NP types (generally in the form of a deviation from a language's prevalent paradigm pattern found only in some minoritarian NP type). The NP type most frequently affected by asymmetrical case marking is the personal pronoun, but other pronoun types, as well as semantically defined subclasses of full nouns (e.g. nouns referring to humans, or personal names), have also been identified as capable of showing deviant case inventories. Classes of nominals based purely on morphological criteria, such as declension classes lacking a common semantic basis or entirely irregular isolated word paradigms, do not constitute NP types, and hence do not manifest asymmetrical case marking.

A typical example of a **case-symmetrical** language is Turkish. All nominals inflect for the same case categories.

(1) Turkish (Kornfilt 1997: 212, 214, 281)

	'book' (SG)	1SG personal pronoun
Nominative	*kitap*	*ben*
Accusative	*kitab-ı*	*ben-i*
Genitive	*kitab-ın*	*ben-im*
Dative	*kitab-a*	*ban-a*
Locative	*kitap-ta*	*ben-de*
Ablative	*kitap-tan*	*ben-den*
Instrumental	*kitap-la*	*ben-im-le*

A language is **additive-quantitatively case asymmetrical** when a particular minority NP type inflects for more categories than the general noun paradigm. The most straightforward example is English, with its objective case represented only in the (majority of the) personal pronouns and in the interrogative-relative pronoun *who*, but lacking from the other nominals.

(2)

	'house' (SG)	1SG personal pronoun
direct objective	*house*	*I* / *me*
genitive	*house's*	*my*

Case marking on full nouns in English is limited to the "Saxon" genitive (clitic case marking; see Chapter 49). An additive-quantitatively asymmetrical language without any case on full nouns is Kobon (Trans-New Guinea; Papua New Guinea).

(3) Kobon (Davies 1981: 154)

	'peanut'	1SG personal pronoun
direct objective	*gaisam*	*yad* / *ip*

Kobon also serves to explain the numerical difference between the languages exhibiting the feature value "no morphological case" in the present chapter and Chapter 49: languages like Kobon qualify as caseless in Chapter 49, since they lack **generally applicable** case, but count as an instance of feature value three in the present chapter because they do have **pronominal** case.

Case asymmetry is equally possible in languages with rich case inventories. An example in point is Kolyma Yukaghir (eastern Siberia), which has nine distinct cases on full nouns, among them an accusative marked by *-gele*. The 1/2SG/PL personal pronouns, however, have, in addition to the regular accusative, a so-called pronominal accusative in *-ul*, which is used when the subject is 1/2SG/PL (full nouns take the nominative under these conditions; Maslova 2003: 4, 94, 234). Finally, additive-quantitative asymmetry may also obtain when two or more distinct case functions sharing one formal expression in the general case inventory have separate expressions in a particular NP type. This is the situation in Araona (Tacanan; Bolivia), which has a marker *-a* ~ *-ja* covering both ergative and genitive functions on full nouns, but distinct case forms for these functions in several personal pronouns (e.g. 1SG ergative *yama*, genitive *quima*; Pitman 1980: 14, 75, 82).

Subtractive-quantitative asymmetrical case marking is the converse of additive-quantitative asymmetry, i.e. a particular NP type shows fewer categories than the general nominal case inventory. This is exemplified by Chittagong Bengali, which has a distinct ergative for past-tense transitive agents with full nouns but not with pronouns (animate transitive patients apparently take the objective case in all tenses).

(4) Bengali (Učida 1970: 31, 38) (partial paradigm)

	'sister' (SG)	2SG personal pronoun
nominative	*bòin*	*tui*
ergative	*bòinje*	
objective	*bòinjɔrè*	*torè*

Several languages are classified as subtractive-quantitatively asymmetrical only under the stipulation that possessive marking and other adnominal word forms which, in addition to their relational function, encode gender/number agreement with their head nouns not be considered case (but rather adjectives). In Fur (Nilo-Saharan; Sudan), for instance, there is a non-agreeing genitive morpheme *-iŋ* on full nouns which is a true case marker, while personal pronouns in possessor function are adjectives because they agree in number with the possessum: 1SG *duiŋ* (singular possessum), *kuiŋ* (plural possessum) (Jakobi 1990: 97–8, 122–3). Despite the fact that Fur pronouns have some kind of morphological equivalent to the genitive case found with nouns, this language is for formal reasons counted as an instance of subtractive-quantitative asymmetry.

Yimas (Lower Sepik-Ramu; Papua New Guinea) shows subtractive-quantitative asymmetry with respect to toponyms

referring to topographic features (i.e. not to settlements). This language has a general oblique case suffix *-n* ~ *-nan*, which marks various peripheral functions, among them the locative. However, names of topographic features do not take this suffix when occurring in locative function, but remain zero-marked instead (Foley 1991: 165, 170–1). The Yimas example shows that case asymmetry is by no means restricted to pronouns: any semantically coherent class of nominals can constitute a relevant NP type.

While quantitative asymmetry implies a difference in the number of case categories between the general case system and the asymmetry-affected NP types, **qualitative asymmetry** obtains when the boundaries between the functional domains of cases are drawn so differently across the NP types that it is not possible to equate the categories found in the two inflectional systems with each other and to refer to them with the same case labels. Typically, different sets of markers are employed for these categories. In reality, this type is restricted to core-argument alignment splits (splits following the lines of NP-type distinctions). In Pitjantjatjara (Pama-Nyungan; South Australia) full nouns follow ergative-absolutive alignment, while all personal pronouns show nominative-accusative alignment.

(5) Pitjantjatjara (Bowe 1990: 9–19) (partial paradigm)

	'man' (SG)		1SG personal pronoun
Absolutive	*wati*	Nominative	*ngayu-lu*
Ergative	*wati-ngku*	Accusative	*ngayu-nya*
Purposive/ Genitive	*wati-ku*	Purposive/ Genitive	*ngayu-ku*
Locative/ Instrumental	*wati-ngka*	Locative/ Instrumental	*ngayu-la*

Case asymmetry bears superficial resemblance to **case syncretism**, the merger in formal expression of two or more case categories (see also Chapter 28). However, case syncretism implies that the categories affected by the merger are underlyingly still intact and present in the inflectional paradigm, whereas case asymmetry entails the total lack of a paradigmatic category, either in the general case system or in the asymmetry-affected NP type. Syncretic paradigms always have fewer overt case distinctions than the general system, which leads to their easy confusion with subtractive-quantitative asymmetry. In a few instances where both interpretations are feasible, the more plausible one has been adopted here. Case syncretism may take place in single word paradigms or (nonsemantically defined) declension classes as well as in semantically defined NP types. Only the latter instances are deemed significant in treating case asymmetry and are therefore represented on the map by a feature value of their own. A clear instance of such an asymmetry-relevant case syncretism is provided by Latvian, where the second-person plural pronoun is affected (a single pronominal person category is considered a relevant NP type).

(6) Latvian (Mathiassen 1997: 43, 65)

	'father' (SG)	1SG personal pronoun	2PL personal pronoun
Nominative	*tēv-s*	*es*	***jūs***
Accusative	*tēv-u*	*mani*	***jūs***
Genitive	*tēv-a*	*manis*	*jūsu*
Dative	*tēv-am*	*man*	*jums*
Locative	*tēv-ā*	*manī*	*jūsos*

Strictly speaking, case asymmetry is a local phenomenon, that is, it applies more to specific paradigms than to languages as a whole. But every language has more than just two NP types with the potential of hosting case asymmetry, and one and the same NP type of a language may be affected in multiple ways by asymmetry-related phenomena. It is therefore logical that combinations of different types of asymmetry within the same language can and do occur. Such combinations have been dealt with on the map by prioritizing them according to the following ranking hierarchy:

qualitative asymmetry > *additive-quantitative asymmetry* > *subtractive-quantitative asymmetry* > *relevant syncretism*

This means that if in a language, for instance, additive- and subtractive-quantitative asymmetry co-occur, the former phenomenon outranks the latter, and the language is coded on the map as additive quantitative.

2 Theoretical issues

The difference between case asymmetry and case syncretism depends crucially on the definition of inflectional paradigm. Case asymmetry implies the nonexistence of certain case categories in some subset of the nominals, while syncretism presupposes the existence of the same categorial distinctions in all NP types. The case inventory of the general system, i.e. the set of categories defining the nominal morphology of the language ("system-defining structure properties" in the sense of Wurzel 1984), is delimited by the number of distinctive form–meaning pairings that occur in the vast majority of the nominal lexemes. In other words, it is **type-frequency** in the lexicon that makes a particular paradigm structure system-defining. Additional categories, represented only in type-infrequent NP types, are not part of the general system. It is morphologically inadequate to posit, for example, an underlying objective case for all nominals in English, which would surface only in a handful of pronouns but syncretize with the direct case in all other instances (cf. Haspelmath 2002: 145).

Languages may or may not allow case categories such as locative or instrumental, whose prototypical hosts are inanimate nouns, to apply to animates or humans. For the purposes of this chapter, these categories are treated as **semantically incompatible**. Personal pronouns, which are generally coreferential with human nouns, are then excluded from these cases too, and are thus not counted as case asymmetrical.

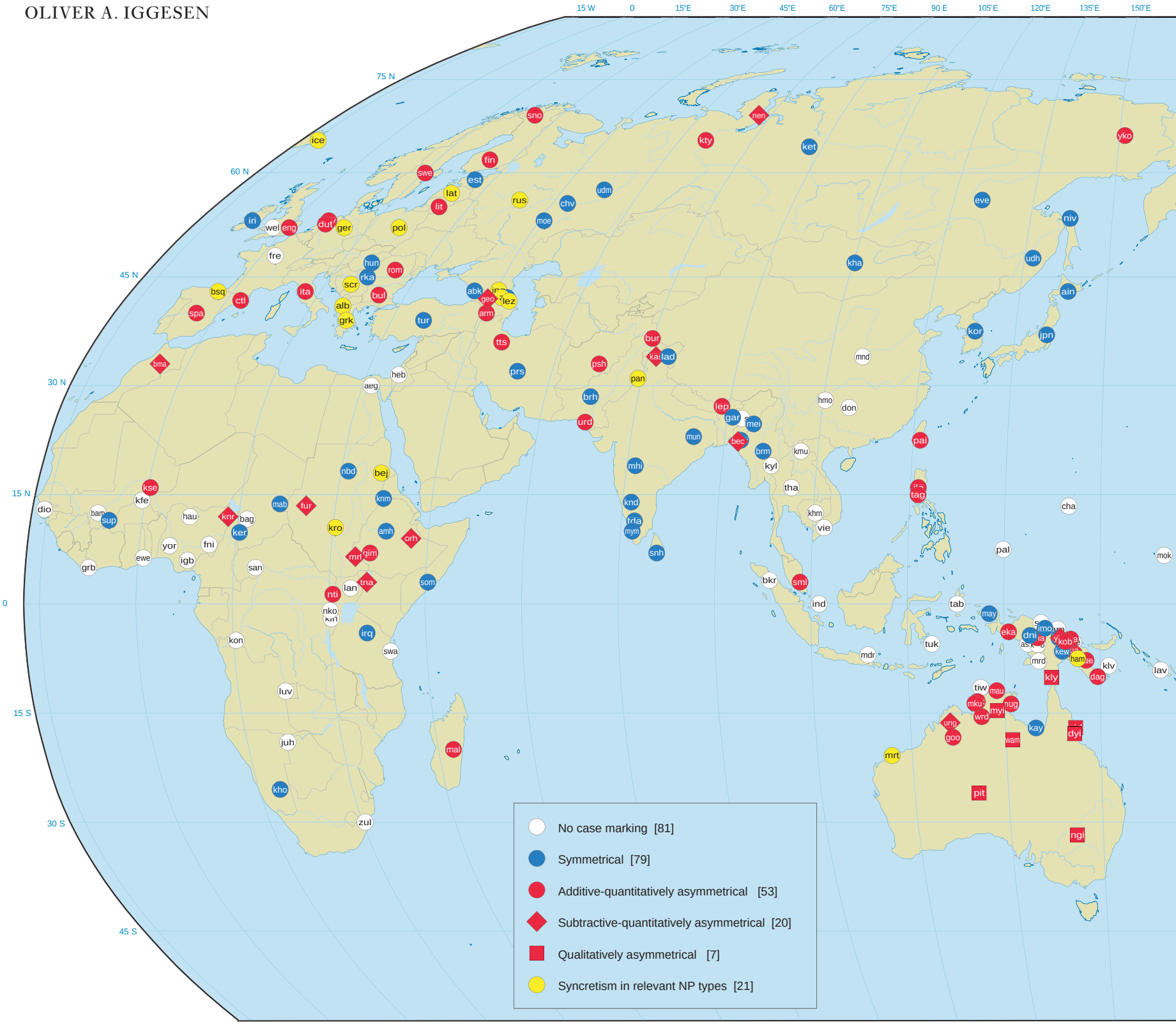
Only **free-form pronouns** are considered valid NP types susceptible to case asymmetry, while **clitic pronouns** (and bound pronominals) have been consistently ignored. Clitics are in general closely associated with the verb phrase, and they are reduced in their status as autonomous nominal constituents. They tend to inflect only for core-argument cases, which would lead to a trivial, mechanistic statement of case asymmetry for practically all languages that have core-constituent clitics and some additional peripheral case in the general system. They also tend to involve too much allomorphy to build neat case paradigms.

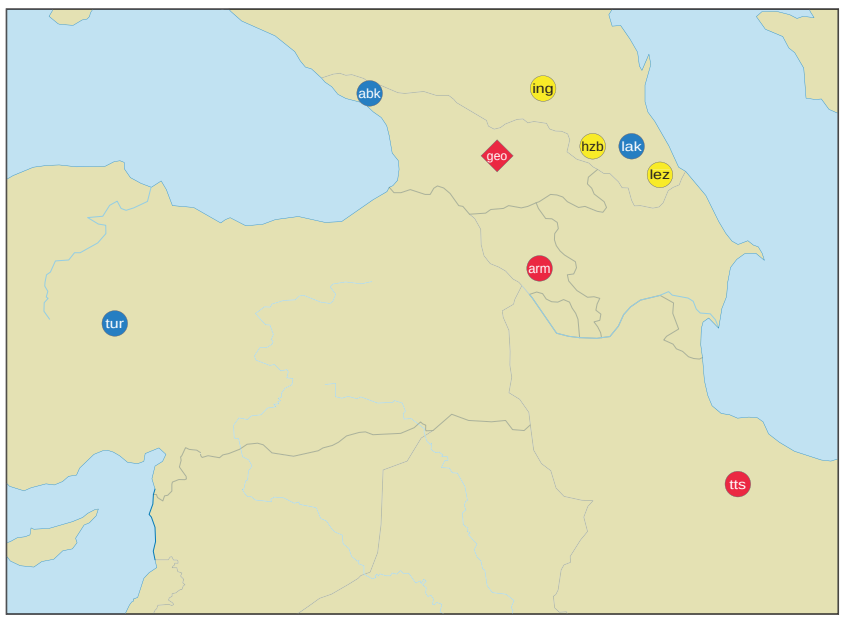
3 Geographical distribution

The map does not show many clear areal patterns; this was perhaps to be expected, given the abstractness of the property and the impossibility of making more fine-grained distinctions (such as representing whether grammatical, peripheral, or adnominal cases are affected by the asymmetry). There is a tendency for asymmetry to be absent in the agglutinative case-marking languages of Eurasia, while significant syncretism is widespread among the more conservative fusional Indo-European languages, and also in the Caucasus. Clusters of additive-quantitative asymmetry are found in Europe, western North America, New Guinea, northern Australia, and the central belt of sub-Saharan Africa. Other areas where additive-quantitative asymmetry is quite common are northern Asia, the Himalayas, and the Amazon. Also most Philippine-type languages show additive-quantitative asymmetry. Subtractive-quantitative asymmetry is found scattered across the map, while qualitative asymmetry is restricted to Australia (among the sample languages; it has also been reported for some Panoan languages; cf. Valenzuela 2000).

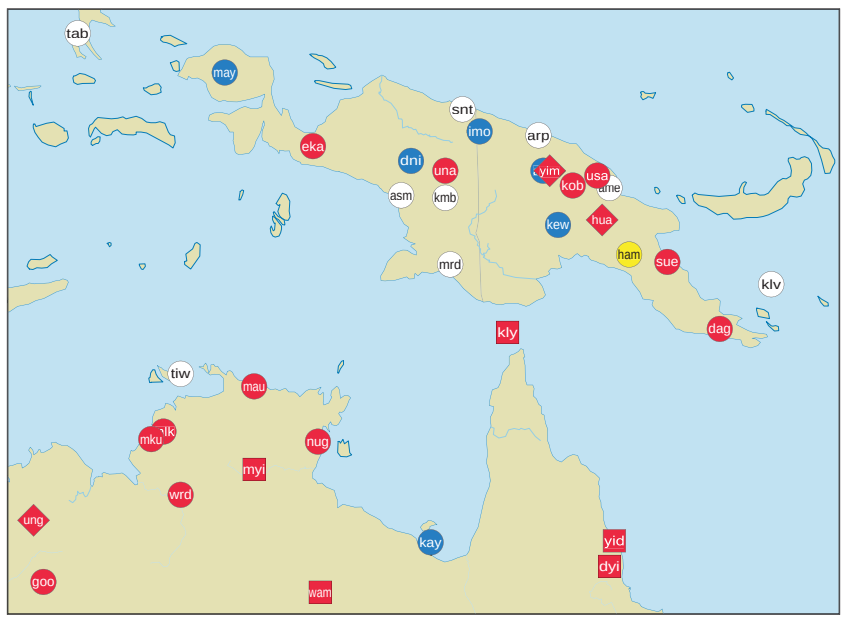
50 Asymmetrical Case Marking

OLIVER A. IGGESEN

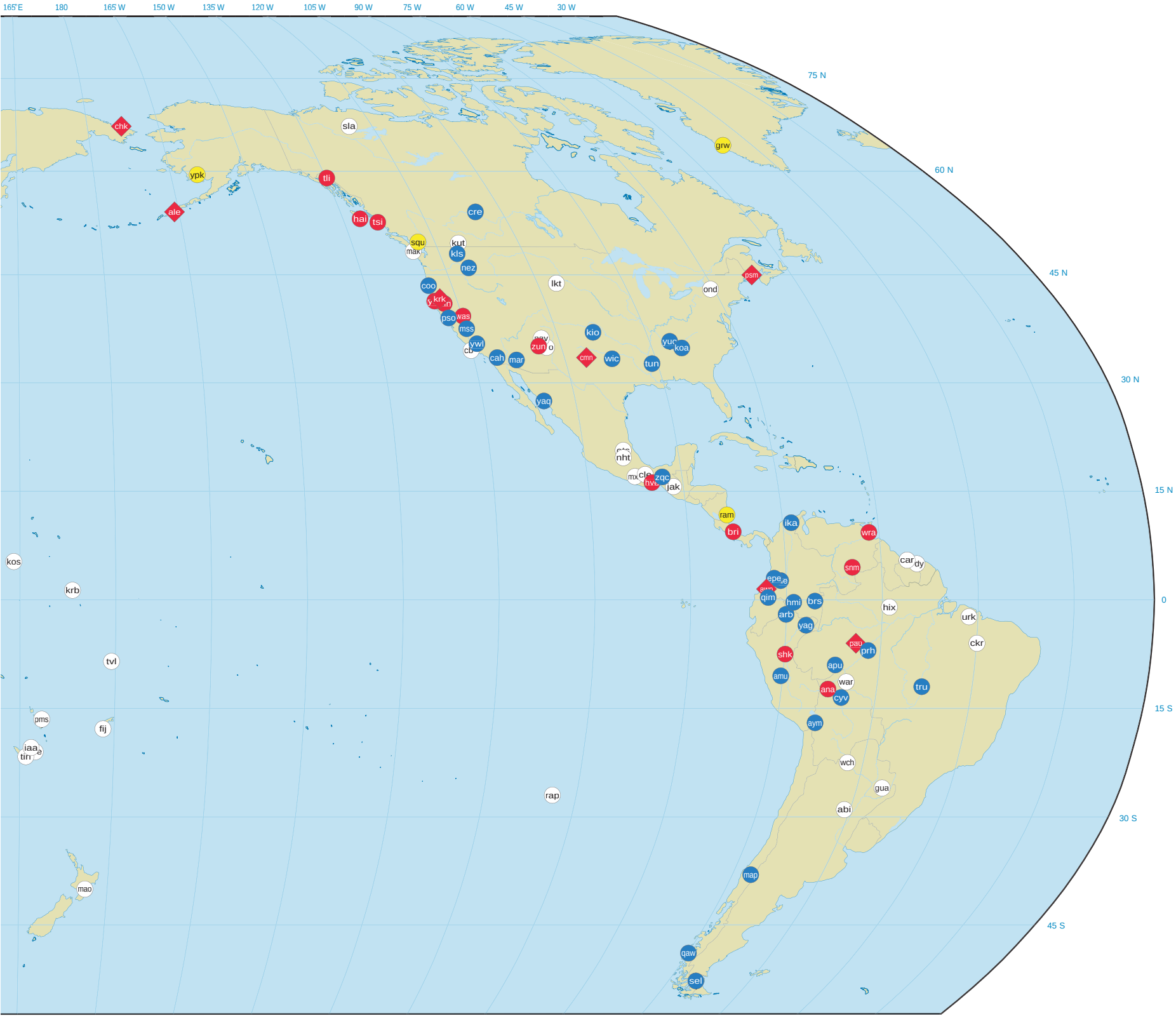


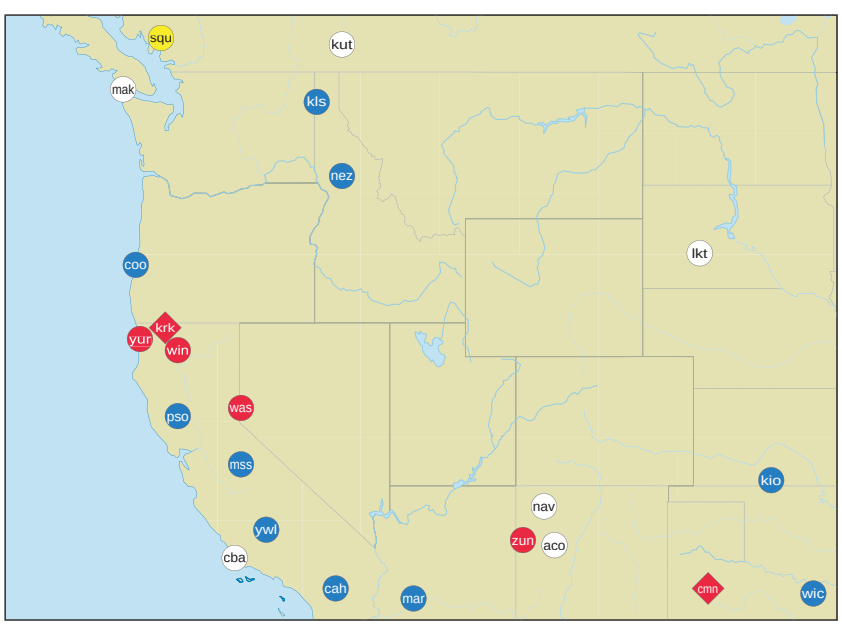


Caucasus and adjacent areas

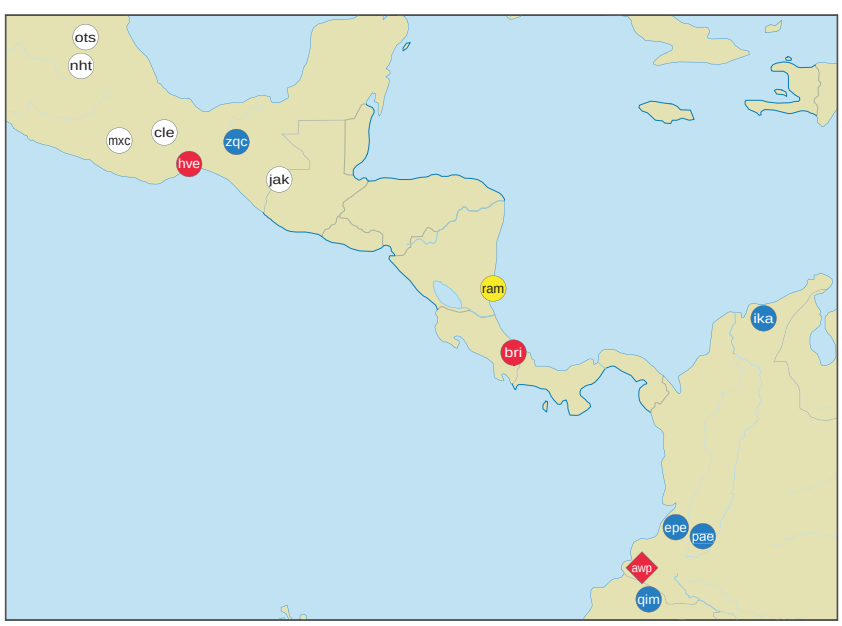


New Guinea and northern Australia





western United States



southern Mexico to Columbia

51 Position of Case Affixes

MATTHEW S. DRYER

1 Defining the values

This map shows the position of case affixes on nouns, especially whether they are **prefixes**, as in the example in (1a) from Krongo (Kadugli; Sudan), or **suffixes**, as in (1b) from Tsova-Tush (Nakh-Daghestanian; Georgia).

(1) a. Krongo (Reh 1985: 151)
àpá-ŋ *ìʔìŋ* *á-kʊ̀ʊ̀fì*
M.PERF.hit-TR 3SG M INSTR-baton
'He hit him with the baton.'

b. Tsova-Tush (Holisky and Gagua 1994: 194)
Lek'-i-v *že* *Dik'-e-r*
Lek-PL-ERG sheep take.away-AOR-PST
'The Leki would take away sheep.'

1.	Case suffixes	431
2.	Case prefixes	35
3.	Case coded by tone	4
4.	Case coded by changes within noun stem	2
5.	Mixed morphological case strategies with none primary	8
6.	Postpositional clitics	95
7.	Prepositional clitics	15
8.	Inpositional clitics	6
9.	Neither case affixes nor adpositional clitics	338
	total	934

A morpheme is considered a case affix for the purposes of this map if it attaches to nouns and signals a grammatical or semantic relation to some other word, most commonly a verb. Some case affixes mark arguments of the verb, such as the ergative case affix in (1b), while others signal the semantic relationship of a noun phrase that is not an argument, as with the instrumental case affix in (1a).

While case prefixes are much less common than case suffixes, they are nevertheless attested. In addition to the example in (1a) are those in (2).

(2) a. Tonga (Bantu; Zambia; Collins 1962: 54)
wakaboola ***a-****Ɉoni*
he.came **with**-John
'He came along with John.'

b. Gurr-goni (Burarran; Australia; Green 1995: 93)
burr-*wupunj* *awurr-bogi-ni*
INSTR-canoe 3AUG.SUBJ-go-PRECONTEMP
'They all went by canoe.'

c. Middle Atlas Berber (Morocco; Penchoen 1973: 54)
iccˇa **u-***ryaz* *aḵsum*
eat.PERF.3SG.M SUBJ-man meat
'The man has eaten the meat.'

d. Prasuni (Nuristani; Afghanistan; Morgenstierne 1949: 220)
ɛsl'ɛk ***tu-****g'ul*
that **in**-country
'in that country'

e. Takelma (Takelman; Oregon; Sapir 1922: 241)
gwel-*danà*
under-rock
'under the rock'

In addition to prefixes and suffixes, there are two other morphological processes which are used to code case, though they are fairly uncommon as primary strategies. One of these involves coding case by **tone**, which is shown on the map as the primary method in four languages: Maba (Maban, Nilo-Saharan; Chad; Tucker and Bryan 1966: 199) plus three Nilotic languages, Shilluk (Sudan; Kohnen 1933: 28), Maasai (Kenya and Tanzania; Tucker and Mpaayei 1955: 175–87), and Nandi (Kenya; Creider and Creider 1989: 41). The examples in (3) from Nandi are both of the form V-NP-NP, with only the tone identifying which NP is subject and which is object. Note that this is not coded by word order: (3a) is VSO while (3b) is VOS.

(3) Nandi (Creider and Creider 1989: 124)
a. *kè:r-éy* *kípe:t* *la:kwé:t*
look.at-IMPF Kipet.SUBJ child NONSUBJ
'Kibet is looking at the child.'

b. *kè:r-éy* *kipe:t* *kípro:no*
look.at-IMPF Kibet NONSUBJ Kiprono.SUBJ
'Kiprono is looking at Kibet.'

A second infrequent method for coding case is by **changes within the noun stem**. Two languages are shown on the map that employ this as their primary method for signaling case, both of them Western Nilotic languages spoken in Sudan: Dinka (Nebel 1948: 36) and Nuer (Crazzolara 1933: 29). In both languages, there are genitive and locative case forms that involve changes to the vowels within the syllable nucleus. Stem changes that only involve the initial or final segment(s) of a word are treated here as a form of prefixing or suffixing, respectively. For example, in Nias (Austronesian; Sumatra, Indonesia), there is a process of initial mutation that typically involves changing the initial consonant to a different consonant, though in some instances it involves the addition of a prefixal consonant, as illustrated in (4). Among a variety of uses, the unmutated forms are used as ergative case forms, the mutated as absolutives.

(4) Nias (Brown 2001: 69–70)

unmutated	mutated	
fakhe	*vakhe*	'rice'
tanö	*danö*	'land'
si'o	*zi'o*	'stick'
ete	*nete*	'bridge'

Nias is treated on the map as an instance of case prefixing.

Some languages combine two or more of the above four methods for coding case. Most such languages combine case prefixes with suffixes. In some languages with more than one morphological method of coding case, one method can be considered primary. For example, in Chukchi (Chukchi-Kamchatkan; eastern Siberia, Russia), there are thirteen cases, all involving suffixes and three involving prefixes in addition to the suffixes (Dunn 1999: 101). Since all case forms involve suffixes, while only some involve prefixes, suffixes can be considered the primary method in this language. In other languages, however, there may be no strong reason to treat one method as primary, and these are shown on the map as languages with **mixed morphological case strategies with no strategy primary**. For example, Mangarrayi (Northern Territory, Australia) employs a mixture of prefixes and suffixes: the case forms for *malam* 'man' include nominative *ṇamalam* with a case prefix, ablative *malamgana* with a case suffix, and locative *ṇamalamgan* with both a prefix and a suffix (Merlan 1982: 57). Similarly, Siuslaw (Oregon Coast family; Oregon) has three case affixes, an ergative prefix, a locative suffix, and a genitive suffix (Frachtenberg 1922a: 462, 541, 544); with one prefix and two suffixes, neither is considered primary. A few languages with a mixture of strategies combine one of the minor methods with prefixes or suffixes. For example, Turkana (Nilotic; Kenya and Uganda) employs both prefixes and tone (Dimmendaal 1983: 66–7, 215, 260–5).

The first four values on the map all involve morphological methods for coding case. The next three values involve what are treated here as adpositional clitics and on Map 85 as adpositions. In general, the difference between case affixes and adpositions is that the former are phonologically attached to nouns, while the latter are separate words that combine syntactically with noun phrases. **Adpositional clitics** are intermediate between these two in that, while they attach phonologically to some word, the word they attach to need not be a noun, and which word they attach to is determined syntactically. The example in (5a) illustrates a **prepositional clitic** attaching to the first word in a noun phrase from Cayuvava (isolate; Bolivia), while the example in (5b) illustrates a **postpositional clitic** attaching to the last word in a noun phrase in Ungarinjin (Wororan; northwestern Australia).

(5) a. Cayuvava (Key 1967: 51)
ji=[ka'reeča datï]
OBL=[other place]
'in another place'

b. Ungarinjin (Rumsey 1982: 58)
[dambun budaga]=ṛa
[camp their]=LOC
'at their camp'

Because such adpositional clitics are very often described as case affixes (especially when they are postpositional) and since in the simplest examples they do attach to the noun, they are shown on this map. There are probably many more languages with prepositional clitics than those shown on the map. An unusual sort of postpositional clitic is used in Somali (Cushitic; Somalia) for marking subject: it is realized by a change in tone on the last syllable of the noun phrase (Saeed 1993: 142). A few languages have **inpositional clitics** which appear inside the noun phrase, most commonly attaching to the end of the first word in the noun phrase, as in the example in (6) from Yawuru (Nyulnyulan; Western Australia).

(6) Yawuru (Hosokawa 1991: 36)
*[kayukayu=**ni** buru]*
[soft=ERG sand]
i-na-nya-rn-dyarra-yirr mudiga
3-TR-catch-IMPF-1AUG.DAT-PL motorcar
'The soft sands caught our car.'

If a language has both case affixes and adpositional clitics, the map shows the type of case affix rather than the type of adpositional clitic. For example, Russian has both prepositional clitics and case suffixes, but is shown on the map as having case suffixes.

The final value shown on the map is for languages which do not employ case affixes or adpositional clitics. This type is somewhat underrepresented on the map relative to its apparent frequency among the world's languages.

The criteria assumed here for case affixes lead to the inclusion of morphemes that might not be considered case affixes by some linguists. For example, Welsh has a system of initial mutations on nouns, which in many environments has nothing to do with case; but the system is such that when object noun phrases beginning with nouns occur in immediately postverbal position or after the subject, they occur in mutated (lenited) form, while postverbal subjects do not (Williams 1980: 79). In (7), the object *fraw* is the lenited form of *braw* 'fright'.

(7) Welsh (Williams 1980: 79)
Cafodd Ifan fraw.
get.PST.3SG Ifan OBJ.fright
'Ifan got a fright.'

These mutations are thus considered a form of case prefix for the purposes of this map. Also, many Bantu languages have locative prefixes which formally belong to the set of noun class prefixes but which serve the same function as clear instances of locative case affixes; hence these are also treated here as instances of case prefixes, as illustrated by the example in (8) from Zulu (South Africa).

(8) Zulu (Ziervogel et al. 1981: 38)
Ni-phuma e-Goli na?
2PL-come LOC-Johannesburg Q
'Do you come from Johannesburg?'

In fact, in some Bantu languages, such locative prefixes co-occur with a noun class prefix, e.g. Hunde (Democratic Republic of the Congo; Mateene 1992: 70).

2 Geographical distribution

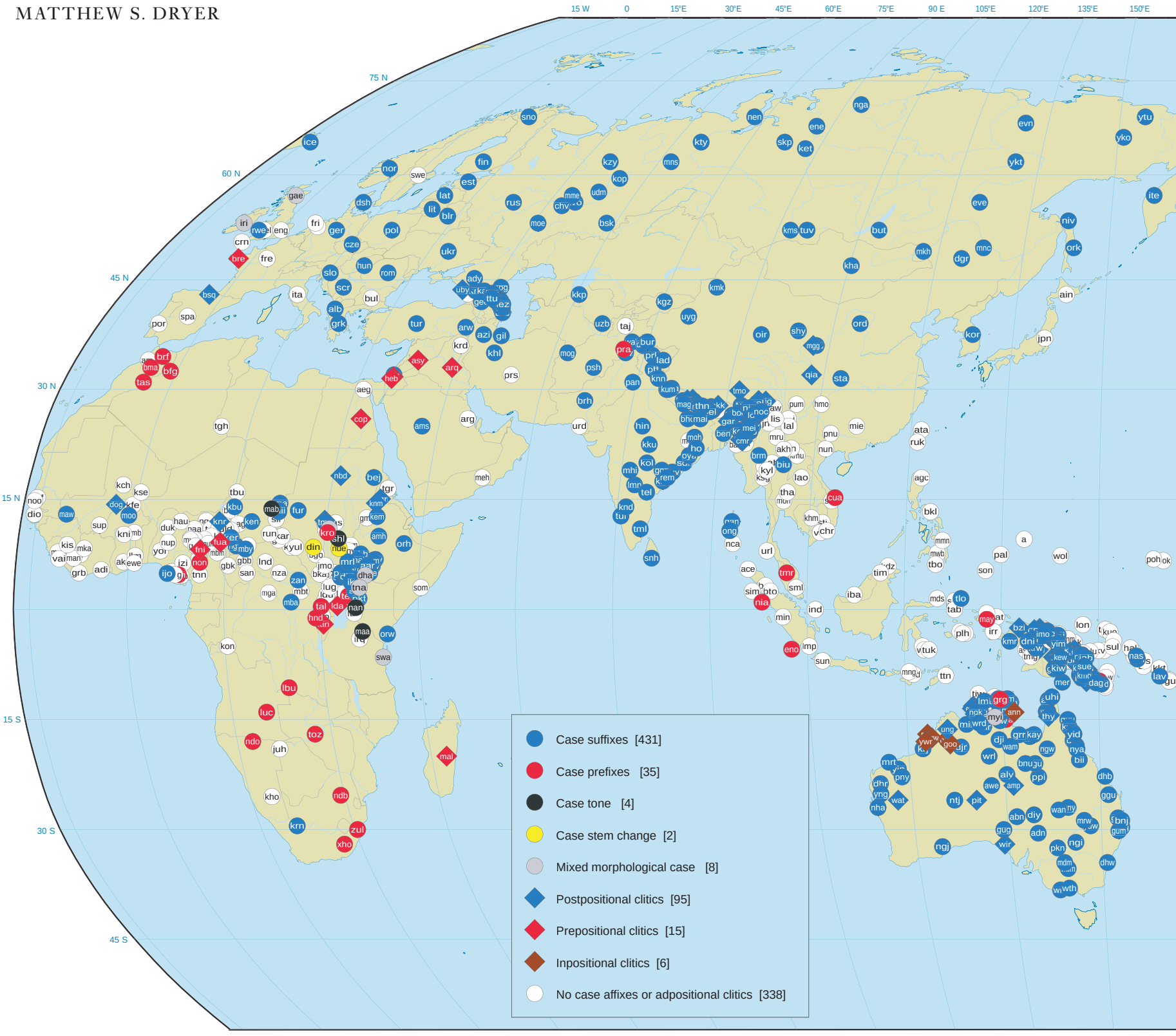
The map shows that case suffixes are overwhelmingly more common cross-linguistically than case prefixes. Apart from isolated instances, the map shows a few areas where more than one language with case prefixes occurs, some of these areas involving only languages which are closely related to each other genealogically, others more heterogeneous. In Africa, case prefixes occur among Bantu languages in the south and in Berber languages in the north. There is a scattering of instances in the area around Sumatra in Indonesia. They are found in Salishan languages in the Pacific North-West and in a scattering of languages in Mesoamerica. Both case by tone and case by stem change as primary strategies are restricted to Africa in the sample of languages examined.

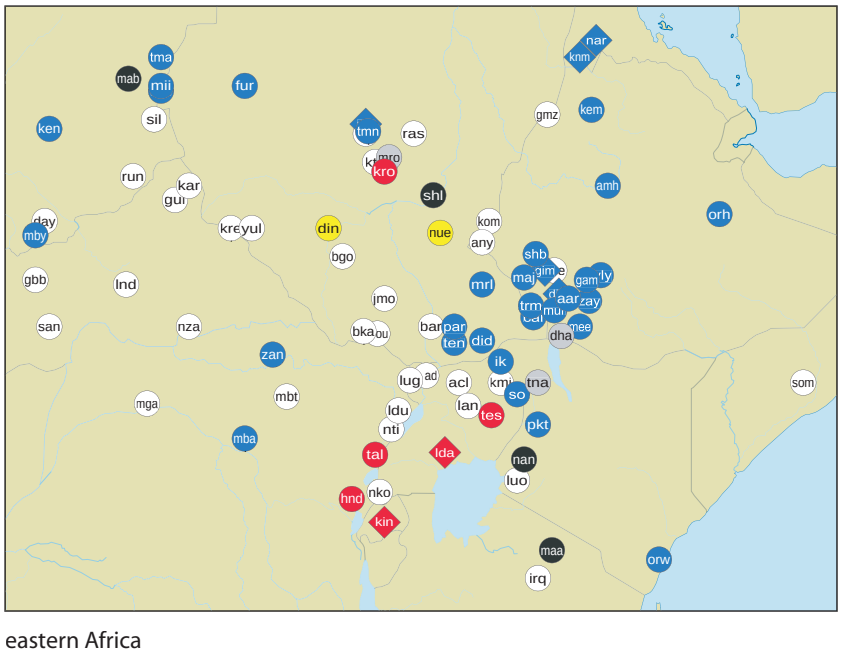
Languages with postpositional clitics tend to occur in areas in which case suffixes are common, such as Australia, North America, and South America. Analogously, prepositional clitics seem to often occur in areas close to where case prefixes are found, but both of these types are sufficiently uncommon that no firm conclusions are possible. The six languages with inpositional clitics are all spoken in Australia, where case suffixes and postpositional clitics are common.

Languages without either case affixes or adpositional clitics are widely distributed. They are particularly common in an area stretching from South-East Asia through the Pacific. Except in eastern Africa, they are the dominant type throughout Africa, though there are many languages of other types. In many areas they are widely scattered among languages that do have case affixes or adpositional clitics, such as the Americas. They are especially infrequent in (i) an area covering eastern Europe and most of Asia, except the South-East and Middle East, and (ii) in most of Australia, except in the far north.

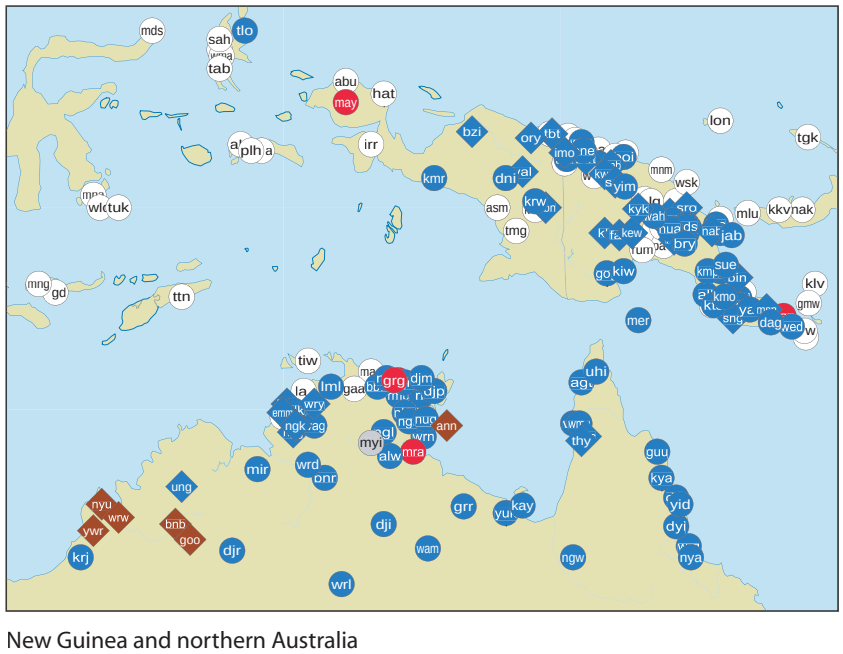
51 Position of Case Affixes

MATTHEW S. DRYER

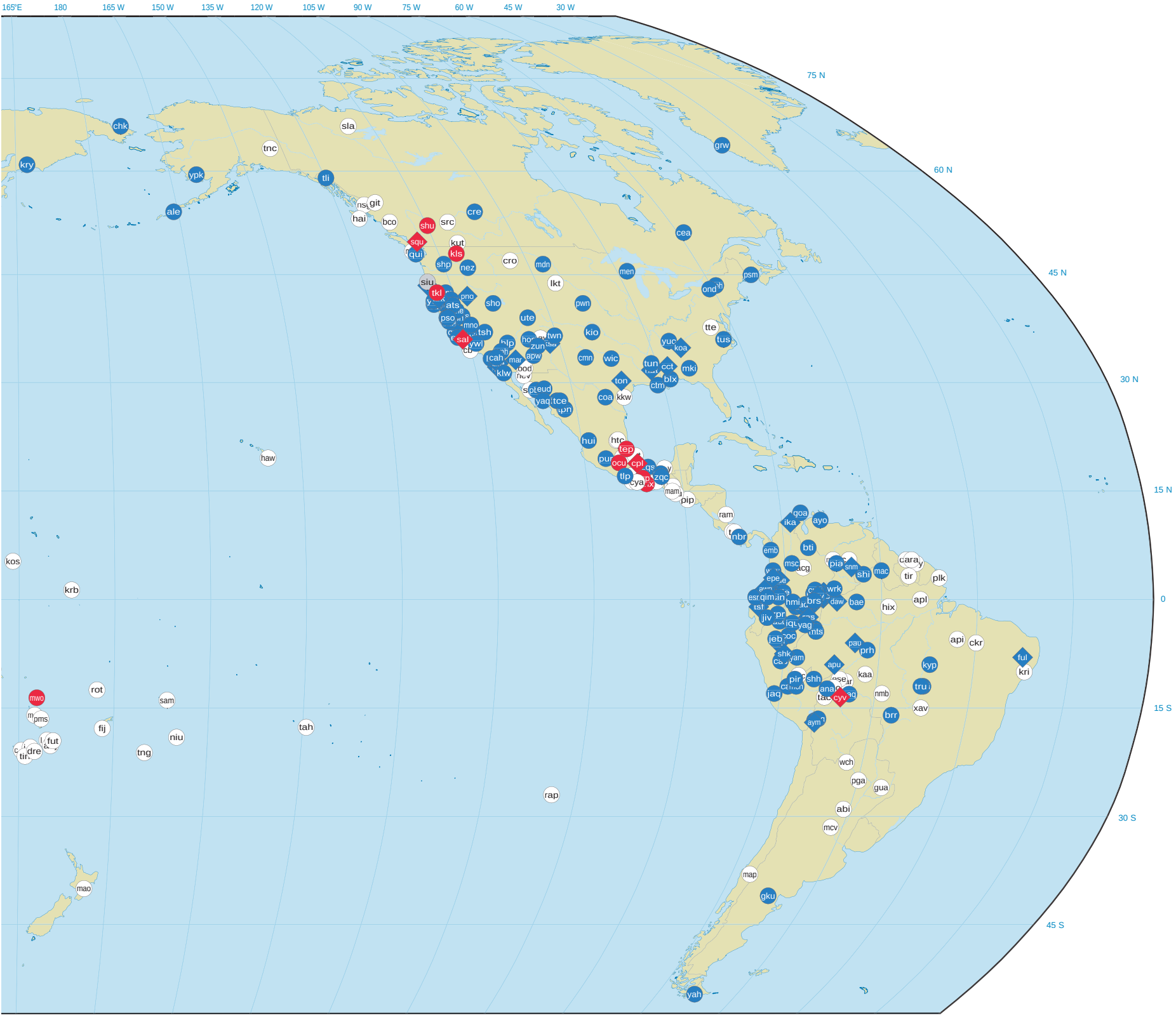


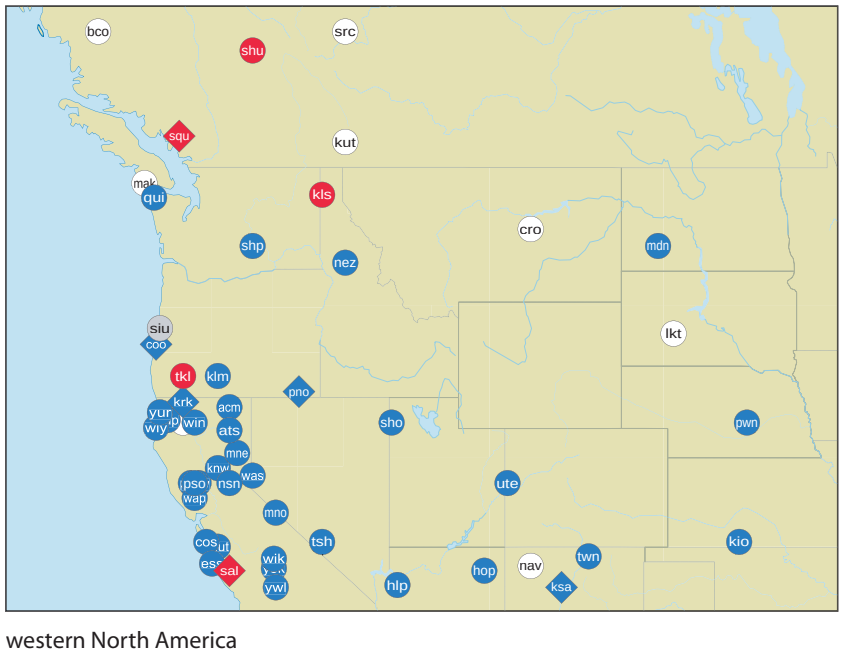


eastern Africa

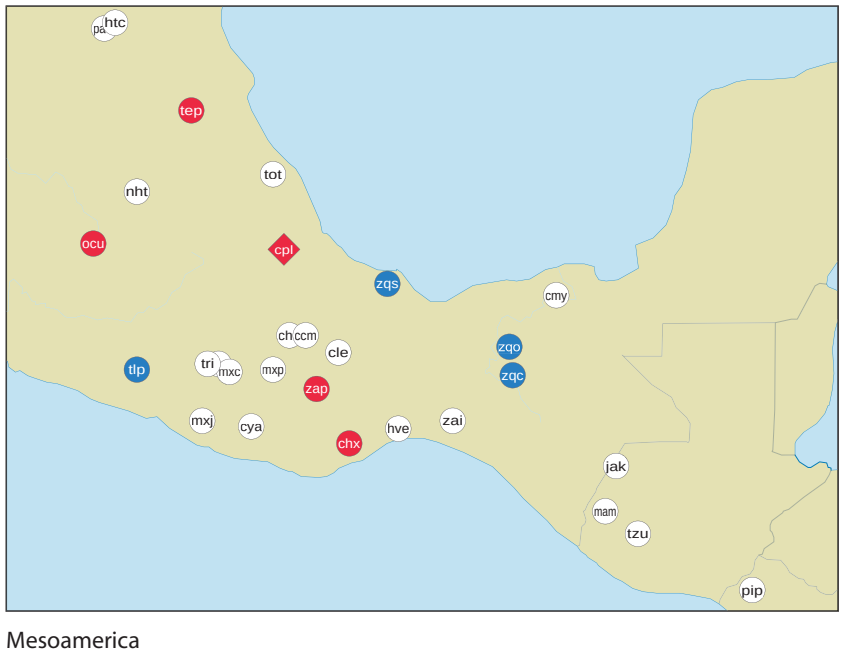


New Guinea and northern Australia





western North America



Mesoamerica

52 Comitatives and Instrumentals

THOMAS STOLZ, CORNELIA STROH, AND AINA URDZE

1 Defining the values

The present chapter provides an overview of the relations that hold between **comitatives** and **instrumentals** in the languages of the world. *Comitative* and *instrumental* are labels for grammaticalized semantic relations between participants of an event. For the purposes of coding such relations, a wide variety of relator morphemes is used, such as case affixes and adpositions. For convenience, all of the relevant relators are glossed by WITH in the examples below. For the present purpose, we have looked only at the most grammaticalized overt morphological expressions. Stylistic and other variation has been discounted for obvious reasons. Moreover, we restrict this presentation to prototypical instances of comitatives and instrumentals.

Prototypically, a comitative relator morpheme is employed to encode accompaniment. An example is the Inga (Quechuan; Colombia) suffix *-hua(n)* in (1a), where the third-person plural subject is the accompanee and *alcalde* 'mayor' is the companion. An instrumental relator morpheme normally marks a noun phrase as the instrument used by an agent in a given situation to carry out the action designated by the lexical verb. An example is (1b), where the same Inga case suffix connects the third-person plural agent with the instrument *caspi* 'stick' used to carry out the action of measuring depths.

(1) Inga (Levinsohn 1976: 124–5)

a. Accompaniment:
alcalde-huan-ta-si *pueblo-ma* *samuncuna*
mayor-WITH-even-QUO town-ALL come.3PL
'They came to town together with the mayor.'

b. Instrument:
caspi-hua-si *tanteancuna* *ñayapa* *junda*
stick-WITH-QUO measure.3PL how full
'It was so full, they were measuring it with a tree-trunk.'

Inga is in accordance with the situation found in English insofar as both languages make use of only one relator—the case suffix *-hua(n)* and the preposition *with*, respectively—to encode the two relations under examination. Accompaniment and instrument receive the same coding and, thus, comitative and instrumental are morphologically indistinguishable. This **identity** of coding is rather common in the languages of Europe and also attested elsewhere on the globe (cf. §2). However, comitative-instrumental identity is not the only way for languages to organize the relation between the two categories. In many languages, we encounter the opposite of identity, viz. comitative-instrumental **differentiation**, as well as a mixture of features of identity and differentiation in one and the same language. The three patterns we have been able to identify in earlier publications (see Stolz 1996) are as follows:

1. Identity	76
2. Differentiation	213
3. Mixed	33
total	322

1.1 Identity. This is the pattern valid for English and Inga (cf. (1) above). The relators for comitative and instrumental are identical. The two categories therefore are considered to be part of a syncretistic chain. A non-Indo-European language of Europe characterized by identity is Estonian, cf. (2).

(2) Estonian (Lavotha 1973: 96)

a. Instrument: "Comitative *-ga*"
ma *kirjuta-n* *sule-ga*
I write-1SG pen-WITH
'I am writing with a pen.'

b. Accompaniment: "Comitative *-ga*"
Villem *jaluta-b* *isa-ga*
Villem go.for.a.walk-3SG father-WITH
'Villem is going for a walk with his father.'

The nominal case suffix *–ga* 'with' traditionally bears the label "comitative" (Stolz and Stroh 2001). However, it is used to encode both of the relations in question: it occurs in accompaniment situations on the companion noun (e.g. *isaga* 'with (the) father' in (2b)), and also as the relator marking the instrument (cf. *sulega* 'with (a/the) pen' in (2a)). There are no alternative ways of expressing the two relations with an equally high degree of grammaticalization.

1.2 Differentiation. The most frequent solution for the encoding of comitatives and instrumentals follows a pattern maximally different from that represented by identity. Differentiation requires (at least) two different relators for comitative and instrumental, neither of which can replace the other. This type clearly dominates outside of Europe, although there are also some instances in the Old World, e.g. (3).

(3) Finnish (Karlsson 1978: 125, 133)

a. Instrument: "Adessive *-llä*"
Hän *kirjoittaa* *kynä-llä.*
s/he write.3SG pen-WITH
'S/he is writing with a pen.'

b. Accompaniment: "Comitative *-ine*"
Läsnä *oli* *V. V.* *vaimo-ine-en.*
near be.PST.3SG V.V. wife-WITH-POSS.3
'V. V. was present with his wife.'

In Finnish, there are two inflectional cases which divide up the domain of the single relator of languages of the identity type. The so-called adessive serves inter alia the function of marking instruments (e.g. *kynällä* 'with (a/the) pen' in (3a)), whereas the inflectional comitative *-ine-* indicates accompaniment (e.g. *vaimoineen* 'with his wife' in (3b)). (In recent years, the inflectional comitative has been giving way to an alternative construction with the postposition *kanssa* 'with' governing the genitive case. Irrespective of this ongoing change, the pattern of differentiation has been preserved, because the new construction is also used exclusively to encode comitative function.) Note that Finnish and the closely related and immediately neighbouring Estonian opt for two different solutions: where Finnish employs the pattern of differentiation, Estonian is characterized by identity.

1.3 Mixed. There is a third group of languages that combines features of the two foregoing types, which we call simply **mixed**. In contradistinction to the two above-mentioned patterns, languages of this type have two or more different relators, at least one of which is syncretistic in the sense that it encodes comitative as well as instrumental, and one of which is specialized so that it encodes only either comitative or instrumental. This mixture of features is relatively rare among the world's languages. However, there are instances of "mixed" in Europe, e.g. (4).

(4) Hungarian (Bánhidi et al. 1975: 235, 345)

a. Instrument: "Comitative-Instrumental"
Tol-lal *ír-ok*.
pen-WITH write-1SG
'I am writing with a pen.'

b. Accompaniment: "Comitative-Instrumental"
Jan *is* *megjelenik* *barát-já-val*.
Jan also appear.3SG friend-POSS.3SG-WITH
'Jan too shows up with his friend.'

c. Accompaniment: "Associative"
Csónak-ház-ak *sport-és* *játszóter-ek* *vár-ják*
boat-house-PL sport-ADJ playground-PL wait-3PL
család-ostul *gyerek-estül* *az* *ember-ek-et*.
family-WITH child-WITH DEF man-PL-ACC
'Boat-houses and sports grounds are waiting for the people with family and children.'

In Hungarian, two inflectional case suffixes in the noun paradigm partially compete with each other when it comes to encoding comitatives. First, there is the so-called **comitative-instrumental** suffix *-val/-vál* 'with'. This suffix is polysemous because it may either have a comitative reading marking the companion (e.g. *barátjával* 'with his friend' in (4b)), or it may mark an instrument (e.g. *tollal* (< *toll* + *-val*) 'with a pen'). Thus, *-val/-vál* reflects the familiar pattern of identity. In addition, there is the so-called *associative* case marked by *-stul/-stül* 'with'. The functional domain of this suffix is restricted to the comitative proper, i.e. it marks companions (e.g. *családostul* 'with family', *gyerekestül* 'with child' in (4c)). Put differently, *-stul/-stül* has the typical traits of a morpheme showing differentiation. It is important to note that the associative can never be used with an instrumental reading. However, *-stul/-stül* and *-val/-vál* can be substituted for one another if a comitative reading is intended. Note that Hungarian is a member of the same genealogical group as Finnish and Estonian, though more distantly related to them. Notwithstanding these genealogical ties, Hungarian prefers a third solution which is different from both of those found in Finnish and Estonian.

2 Geographical distribution

Our sample is made up of 322 languages. These languages are asymmetrically distributed over the three types identified above. With a total of 213 languages (= 66 per cent), differentiation clearly dominates on a worldwide scale, leaving only 76 languages, i.e. slightly less than a quarter of the sample, as representatives of identity, and only 33 languages (= 11 per cent) displaying mixed structures. However, the three types show markedly different areal preferences. Notwithstanding the fact that identity is a minor solution as compared to differentiation, it is very strong in Europe, where 24 languages (exactly half of the sample languages spoken in Europe) lump together comitative and instrumental. Outside of Europe, the situation changes dramatically. The share of the identity type drops to 30 per cent in Africa, 19 per cent in the Americas, 17 per cent in Asia, and just 8 per cent in the Pacific, whereas the vast majority of non-European languages keep comitative and instrumental formally apart: differentiation characterizes 60 per cent of the languages in Africa, 70 per cent of the Asian languages, 70 per cent of the languages in the Americas, and as much as 86 per cent of the languages of the Pacific. Differentiation dominates everywhere except Europe. On the other hand, the third type—mixed—is always a minority solution, and never reaches beyond the 20 per cent mark. The highest score for mixed is 18 per cent in Europe; this relatively high percentage probably reflects the tendency (see below) of erstwhile members of the differentiation class to acquire the identity pattern via language contact in Europe. Notwithstanding the marked disproportions in the geographical distribution, none of the three types is totally absent from any of the five areas; cf. Table 1. With the exception of Europe, every continent favours one of the patterns—more precisely, differentiation—with far more than 50 per cent.

Table 1 Distribution of types over areas

type	Africa	Americas	Asia	Europe	Pacific	total
identity	20	15	11	24	6	76
differentiation	41	53	44	15	60	213
mixed	7	8	7	8	3	33
total	68	77	62	47	69	322

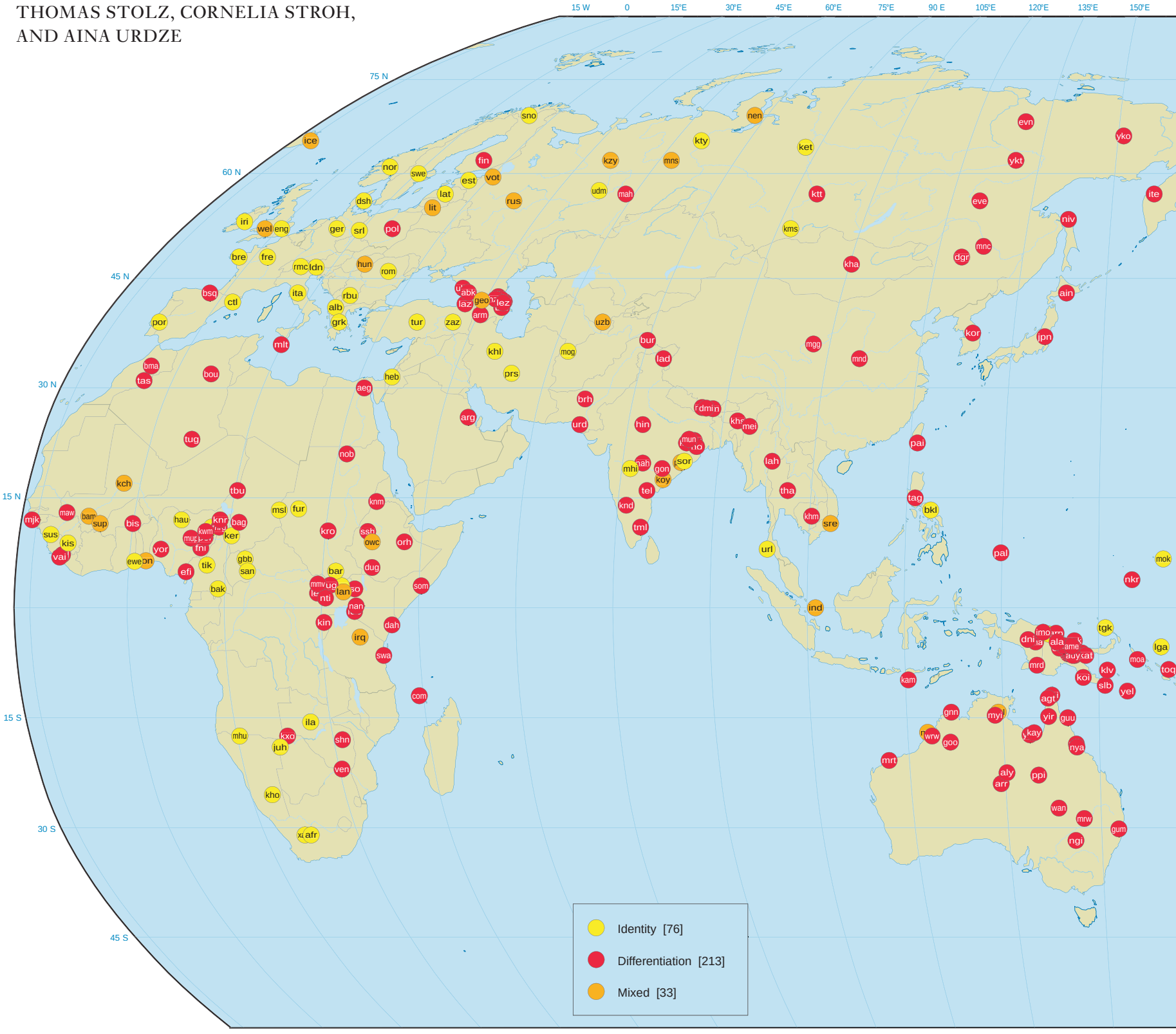
It is interesting to note that even within Europe, where identity has its one real stronghold, the pattern itself seems to be a relatively recent innovation spreading from the Standard Average European languages to their neighbours (Stolz 2001c). The three Finno-Ugric languages from which the above examples are drawn probably differ in their organization of the relation between comitative and instrumental precisely because two of them—Estonian and Hungarian—have been subject to contact influence by Indo-European languages (most prominently German in both cases), where identity is the established pattern.

3 Theoretical implications

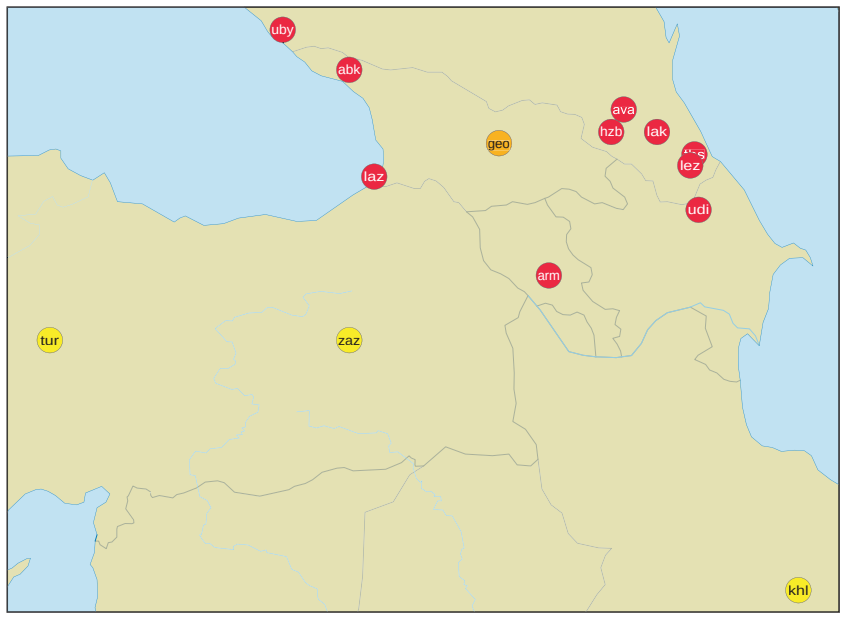
Our findings bear upon current linguistic theory for the following reason. Formerly, the view was widely held that identity is the most frequent pattern worldwide, if not the only possible structural solution for comitatives and instrumentals (Lakoff and Johnson 1980; for other pertinent titles, cf. Schwarz et al. 2001). However, this putative universal of human language structure turns out to be nothing more than an areal preference of languages of Europe—especially those which are counted among the so-called Standard Average European languages. As the examples from the three Finno-Ugric languages suggest, genetic affiliation of languages does not imply identical structural behaviour. This applies to Indo-European languages too, which mostly adhere to identity in Europe whereas their relatives in Asia are divided between differentiation and identity (Stolz 1996). Nevertheless, identity is a relatively frequent pattern, although both the comitative and the instrumental can also combine syncretistically with quite different categories (Stolz 2001a, 2001b, Stassen 2000 and Chapter 63 below).

52 Comitatives and Instrumentals

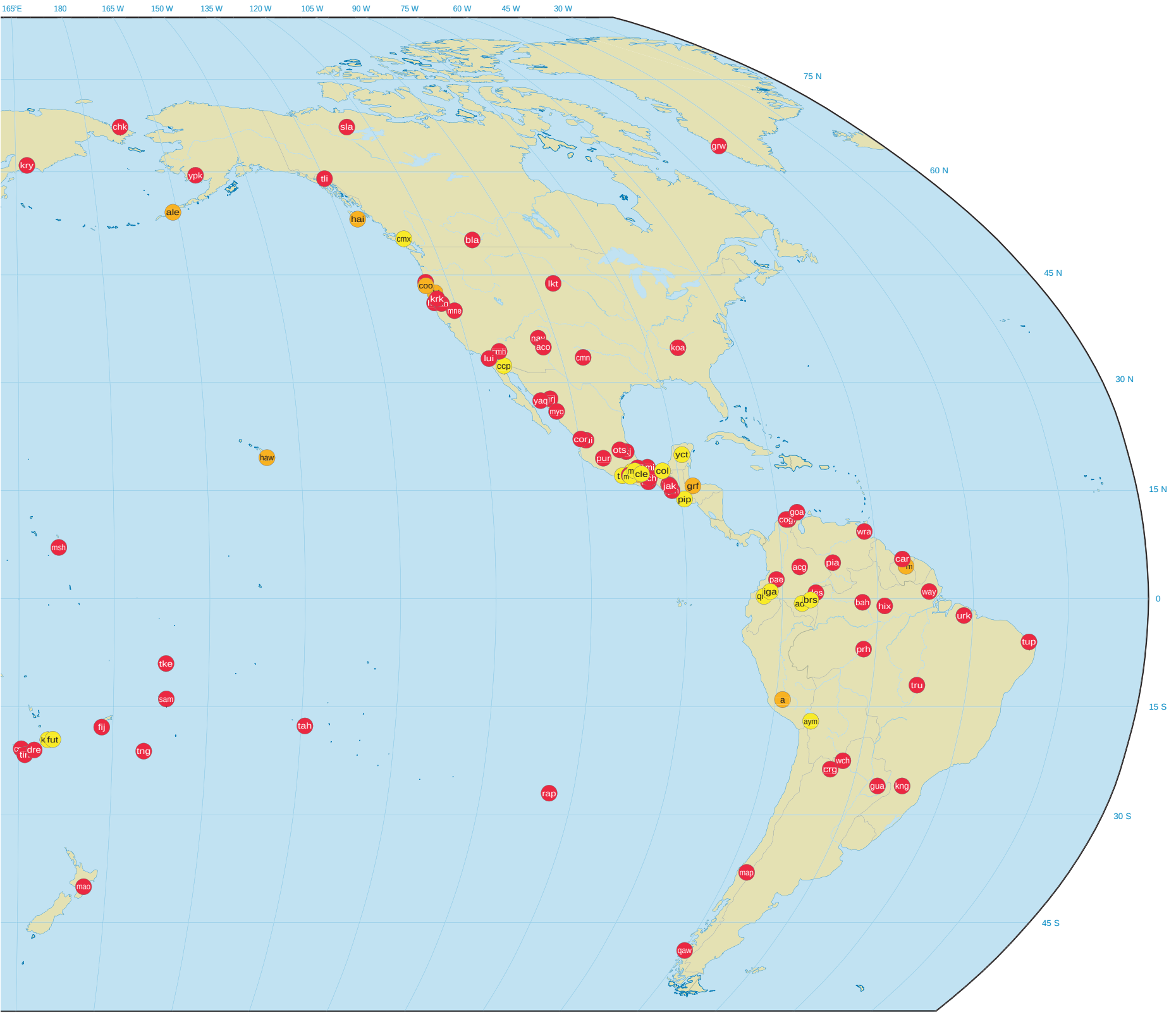
THOMAS STOLZ, CORNELIA STROH,
AND AINA URDZE

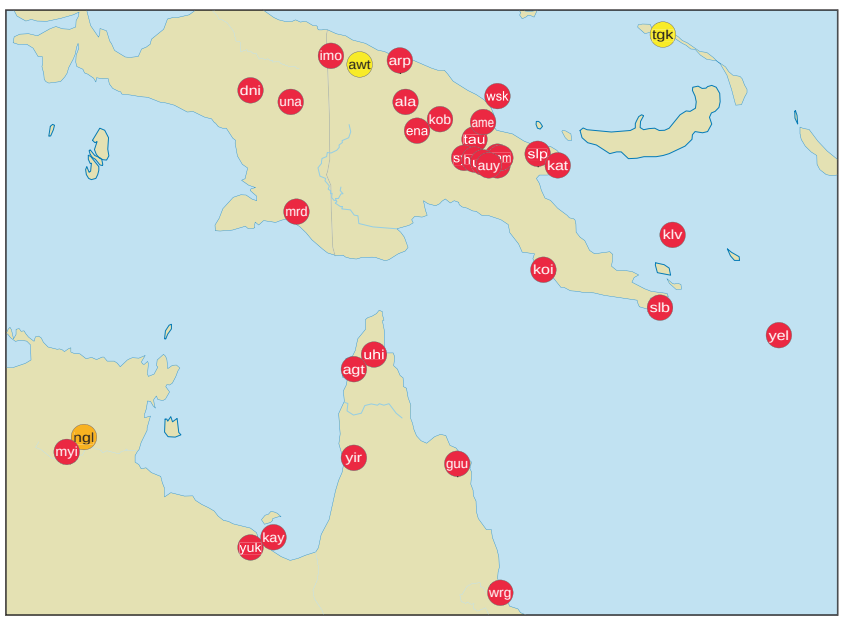


Nigeria and surrounding areas

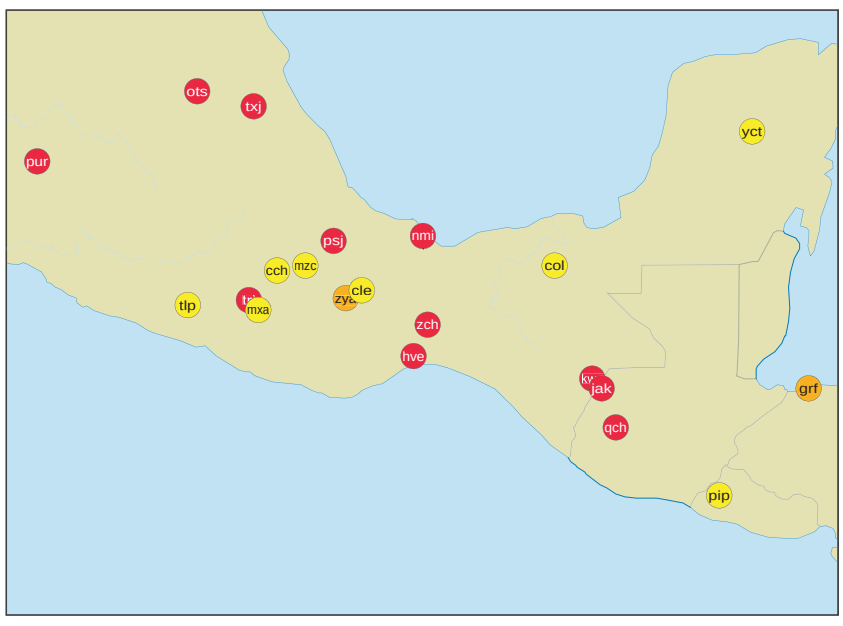


Caucasus and adjacent areas





New Guinea and northern Australia



Mesoamerica

53 Ordinal Numerals

THOMAS STOLZ AND LJUBA N. VESELINOVA

1 Introduction

This article surveys the global distribution of morphological patterns that connect ordinal numerals with cardinal numerals. We associate these two classes of numerals with different functional domains. By **cardinal numerals**, we understand the set of numerals used in attributive quantification of nouns (e.g. 'four days'). **Ordinal numerals** typically identify the position a given member of a set occupies relative to other members of the same set (e.g. 'the fourth day'). The main functions of ordinal numerals thus comprise the identification of ranks within a hierarchy and the identification of the temporal order in a sequence of events or the like. For more information on numerals, see Hurford (1975, 1987).

We look at the formal relationship that obtains between cardinal numerals and their ordinal counterparts, i.e. we study morphological and syntactic (periphrastic) ways of deriving ordinals from cardinals. However, we exclude nonmorphological marking by word order (cf. §2.2). Furthermore, we restrict our investigation to the first decade (i.e. to the values one through ten), because in many languages numeral systems do not go beyond ten, or the use of ordinal numerals is optional with higher numerical values. For the sake of brevity, our examples below list numerals up to value five. The **derivational dependence** of ordinal numerals on cardinal numerals is the cover term for regular morphological derivation according to the synchronically valid rules of a given language, and also for all those cases where the morphological relation is of a purely historical nature, so that the word form of the ordinal numeral reflects obsolete derivational rules of an earlier stage of the same language. Otherwise we consider ordinal numerals **derivationally independent** of their cardinal counterparts, or in other words, **suppletive**.

2 Defining the values

In Veselinova (1997) and Stolz (2001d), we observed that cross-linguistically the morphological relation of cardinal numerals and ordinal numerals is neither random nor homogeneous. We established eight recurrent patterns whose validity has been corroborated by the present sample of 321 languages. These eight types are listed in the feature-value box:

○	1. None: Ordinal numerals do not exist	33
○	2. One, two, three: No distinction of cardinal and ordinal numerals	3
○	3. First, two, three: Cardinal and ordinal numerals are identical except for 'one' and 'first'	12
○	4. One-th, two-th, three-th: Ordinal numerals are derived from cardinal numerals	41
○	5. First/one-th, two-th, three-th: All ordinal numerals are derived from cardinal numerals with two alternatives for 'first', one of which is morphologically independent of 'one'	54
○	6. First, two-th, three-th: Ordinal numerals from 'two' upwards are derived from cardinal numerals, 'first' is suppletive	110
○	7. First, second, three-th: 'First' and a small set of consecutive higher ordinal numerals are suppletive	61
○	8. Various: Other solutions	7
	total	321

2.1 None. In quite a few languages, ordinal numerals do not exist at all, nor are their functions taken over by cardinal numerals. An example is Kobon (Trans-New Guinea; Davies 1989: 207–8), where a system of cardinal numerals is attested, but temporal and local adverbs are used instead to encode the relative order of phenomena: *ñi nöd* (lit. 'child before') 'first/eldest child'.

2.2 One, two, three. What distinguishes this type from the preceding type is the fact that numerals are used to indicate temporal sequence and hierarchical order. Morphologically, there is no way of telling cardinal and ordinal numerals apart, i.e. there is only one multifunctional set of numerals. Sapuan (Bahnaric, Austro-Asiatic; Laos) offers convincing examples of such multifunctional numerals, as in (1).

(1) Sapuan (Jacq and Sidwell 1999: 30–1)

a. *səm* *bar* *lăŋ*
house two house.CLASS
'two houses'

b. *bar* *măŋ* *law*
two language Lao
'my second language is Lao'

In (1a) *bar* is used as the equivalent of the cardinal numeral 'two', whereas in (1b) it corresponds to the ordinal numeral 'second' of the English gloss.

The distinction of cardinal and ordinal functions of numerals via word order alone, as attested in Indonesian, can also be seen as an instantiation of the type "One, two, three" because the morphology of the numeral itself is not affected in any way by the different syntactic orderings, as in (2).

(2) Indonesian (Kwee 1981: 97)

a. *gadis* *ke-tiga*
girl DEF-three
'the third girl'

b. *ke-tiga* *gadis*
DEF-three girl
'the three girls'

If the definite numeral is used in postnominal position (cf. 2a), its function is clearly that of an ordinal. In prenominal position, the numeral can only have a cardinal reading (cf. 2b).

2.3 First, two, three. This type shares the identity of cardinal numerals and ordinal numerals with the foregoing type "One, two, three", but with one exception: there is a separate and distinct expression for 'first' that is derivationally independent of the cardinal 'one'. The inventories of ordinal and cardinal numerals proper therefore consist of only one item each. This is the case in Goajiro (Arawakan; Venezuela), where from two onwards cardinal and ordinal numerals are identical, cf. (3).

(3) Goajiro (Olza Zubiri and Jusayú 1985: 43–5)

Value	cardinal	ordinal
1	*wané*	*pala'jana*
2	*pia'ma*	
3	*apü'nüin*	
4	*piéinchi*	
5	*ja'rrái*	

There is no synchronic or diachronic morphological connection between *wané* 'one' and *pala'jana* 'first'. The first position in the numeral paradigm is subject to suppletion.

2.4 One-th, two-th, three-th. In this type, cardinal and ordinal numerals are formally distinct and none of the ordinal numerals is derivationally independent of its corresponding cardinal partner. All ordinal numerals are regularly derived from matching cardinal numerals. This general rule extends also to 'one' and 'first'. A typical example of "One-th, two-th, three-th" is Hunzib (Nakh-Daghestanian; eastern Caucasus); cf. (4).

(4) Hunzib (van den Berg 1995: 69–70)

Value	cardinal	ordinal
1	*hɔ̃s*	*hɔ̃s.sə*
2	*q'an.u*	*q'an.u.s*
3	*λɑ.nɑ*	*λɑ.nɑ.s*
4	*oq'e.n*	*oq'e.no.s*
5	*λi.no*	*λi.no.s*

The ordinal numerals in this paradigm are derived from the cardinal numerals via suffixation of *-s(ə)*. The phonological shape of this suffix is dependent on the quality of the final segment of the numeral stem. The pair consisting of *hɔ̃s* 'one' and *hɔ̃s.sə* 'first' behaves morphologically exactly like the rest of the numerals. The paradigm is absolutely regular and free of suppletion.

2.5 First/one-th, two-th, three-th. This is a kind of mixture of the two previous types. In principle, regular derivation of ordinal numerals from their corresponding cardinal partners applies on every rank and for every value. However, there are two alternative options for the expression of 'first', one following the regular derivation pattern, the other being suppletive. A good example of this allomorphy is Turkish. In this language, every ordinal numeral is derived from its cardinal counterpart via suffixation of *-(I)ncI* (the vowel quality changes according to the rules of vowel harmony); cf. (5).

(5) Turkish (Christoph Schröder, p. c.)

Value	cardinal	ordinal
1	*bir*	*birinci/ilk*
2	*iki*	*ikinci*
3	*üç*	*üçüncü*
4	*dört*	*dördüncü*
5	*beş*	*beşinci*

Accordingly, we find *birinci* 'first' as a regular derivation of *bir* 'one'. However, alongside this formation there is also the derivationally independent but synonymous form *ilk* 'first'.

2.6 First, two-th, three-th. In this type, all ordinal numerals above 'first' are regularly derived from their cardinal partners, but 'first' is derivationally independent of 'one'. In Kisi (Southern Atlantic; Guinea) we observe that ordinal numerals are regularly derived from cardinals via the suffix *-ndɔ́ɔ́*, except for *tásè* 'first', which is derivationally independent of *pìlɛ̀ɛ́* 'one', cf. (6).

(6) Kisi (Childs 1995: 112–13, 200–1)

Value	cardinal	ordinal
1	*pìlɛ̀ɛ́*	*tásè*
2	*mùúŋ*	*mùúŋndɔ́ɔ́*
3	*ŋgàá*	*ŋgàándɔ́ɔ́*
4	*hìɔ́ɔ́lú*	*hìɔ́ɔ́lúndɔ́ɔ́*
5	*ŋùɛ̀ɛ́nú*	*ŋùɛ̀ɛ́núndɔ́ɔ́*

2.7 First, second, three-th. In this type the derivational independence of ordinals is extended further at least to the second rank. In Estonian, for example, alongside the suppletive expression of 'first', we also find a derivationally independent form for 'second', cf. (7).

(7) Estonian (Lavotha 1973: 52–4)

Value	cardinal	ordinal
1	*üks*	*esimene*
2	*kaks*	*teine*
3	*kolm*	*kolmas*
4	*neli*	*neljas*
5	*viis*	*viies*

The ordinal numeral *esimene* 'first' has no morphological ties to the cardinal numeral *üks* 'one', and the same holds of the relation between the ordinal numeral *teine* 'second' and its cardinal counterpart *kaks* 'two'. The regular derivation of ordinal numerals from cardinal numerals starts with three and is achieved via the suffixation of *-s*.

2.8 Various. The eighth group of languages does not form a proper homogeneous type. It is rather the class of individual solutions and exceptions, including a variety of (contact-induced) mixed systems. A pertinent example of an individual solution is Vietnamese; cf. (8).

(8) Vietnamese (Hoàng Thị Chau 1982: 50)

Value	cardinal	ordinal
1	*một*	*thứ nhất*
2	*hai*	*thứ hai*
3	*ba*	*thứ ba*
4	*bốn*	*thứ tư*
5	*năm*	*thứ năm*

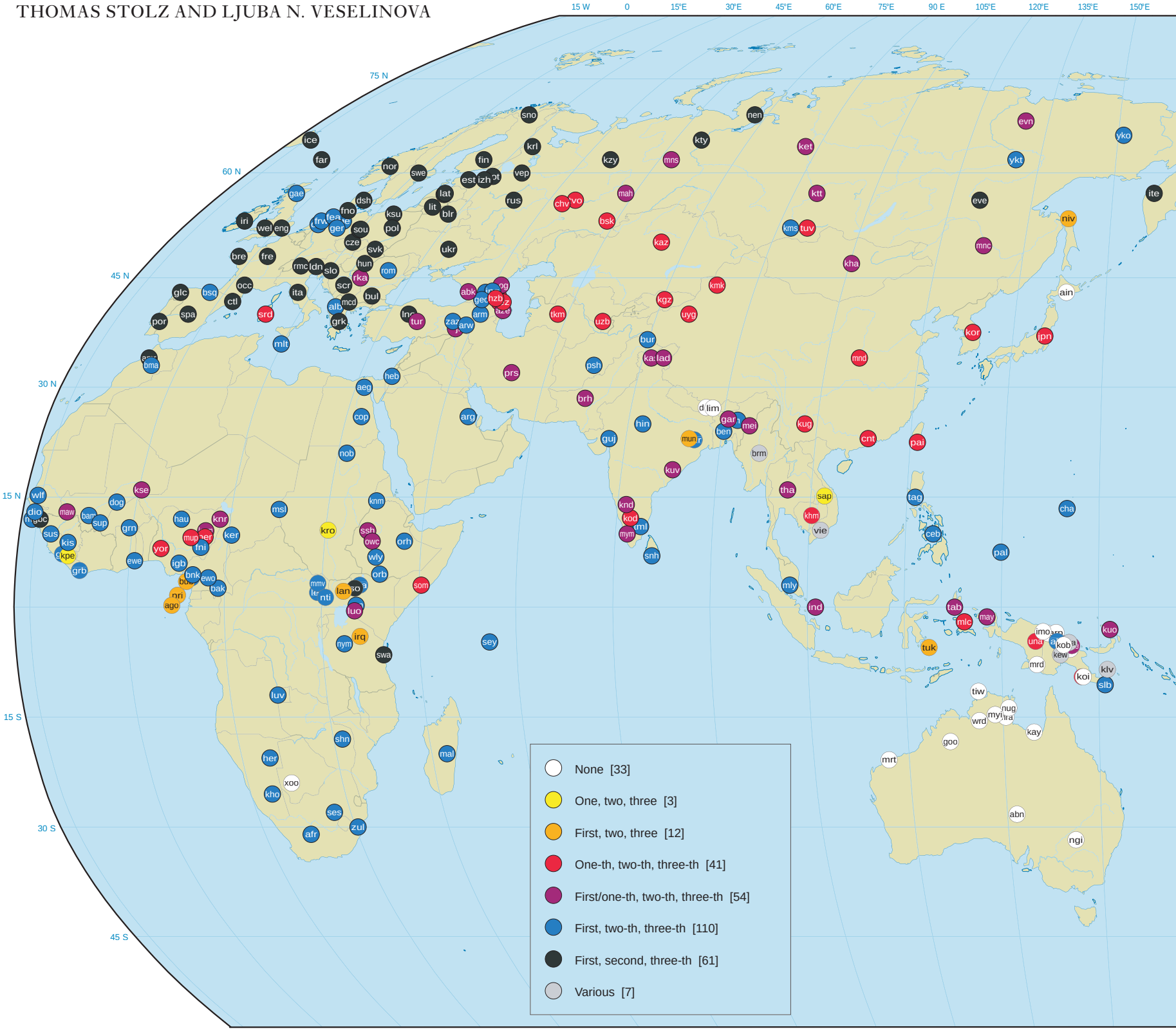
Alongside the suppletive formation of 'first', *thứ nhất*, which is morphologically independent of *một* 'one', there is also *thứ tư* 'fourth', which has no derivational connection to the cardinal numeral *bốn* 'four'. The remaining ordinal numerals are based on their cardinal equivalents. This situation is exceptional insofar as normally suppletion on a value rank $n > 1$ automatically requires suppletion on all value ranks between n and 1.

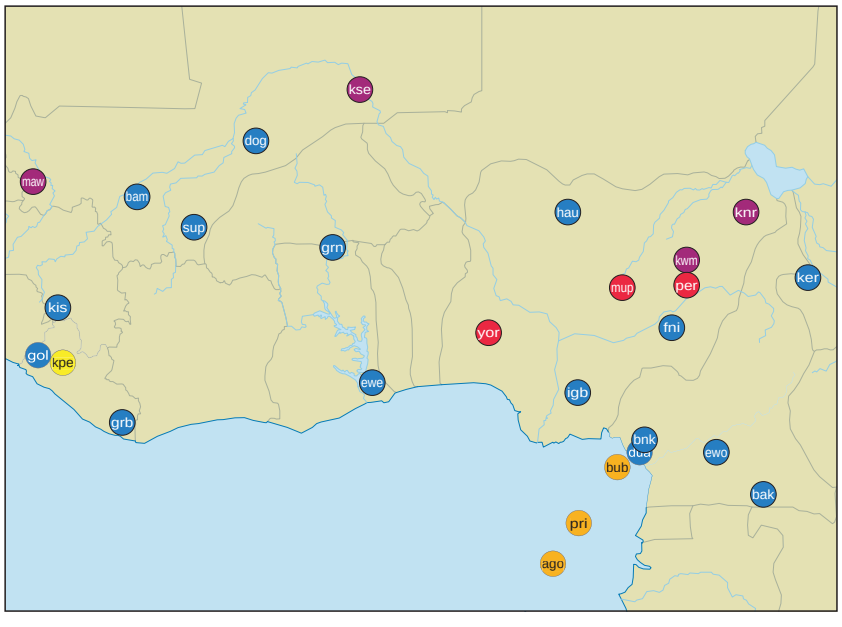
3 Geographical distribution

The geographical distribution of the above types is uneven. This asymmetry is most pronounced with "First, second, three-th": of the 61 languages of this type, 47 (= 76 per cent) are located in Europe. The most frequently attested type, "First, two-th, three-th", comprises 110 languages on all continents. Of these, 44 (= 40 per cent) are spoken in Africa and 23 (= 21 per cent) in the Americas. The types "One-th, two-th, three-th" and "First/one-th, two-th, three-th" are especially frequent in Asia (14 and 18 languages = 34 per cent and 33 per cent respectively), whereas "None" has two strongholds in the Americas and in the Pacific region with 12 and 17 languages (= 36 per cent and 51 per cent respectively). Asia is the only continent where all of the above types are attested at least once. The types "One-th, two-th, three-th", "First/one-th, two-th, three-th", and "First, two-th, three-th" are attested on all continents, the remaining types being absent from at least one region.

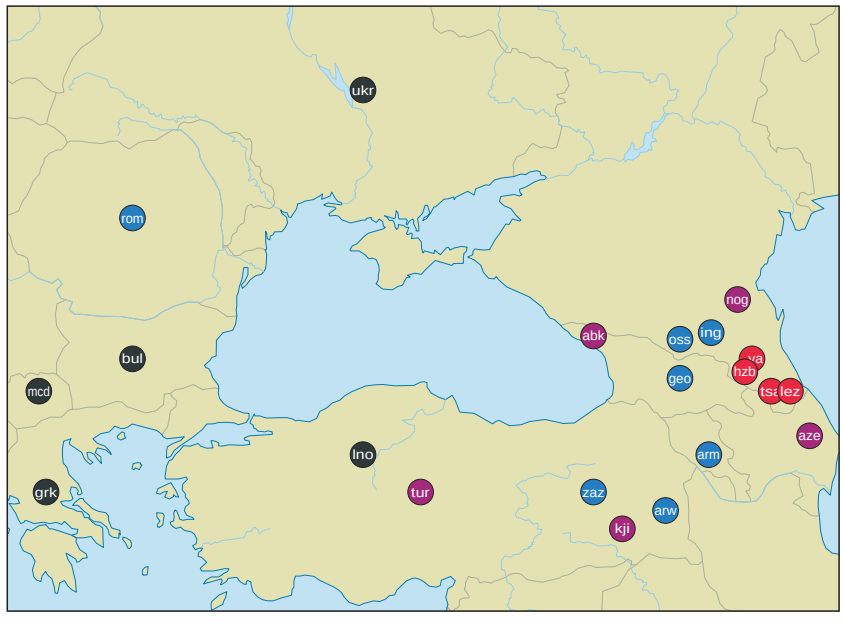
53 Ordinal Numerals

THOMAS STOLZ AND LJUBA N. VESELINOVA

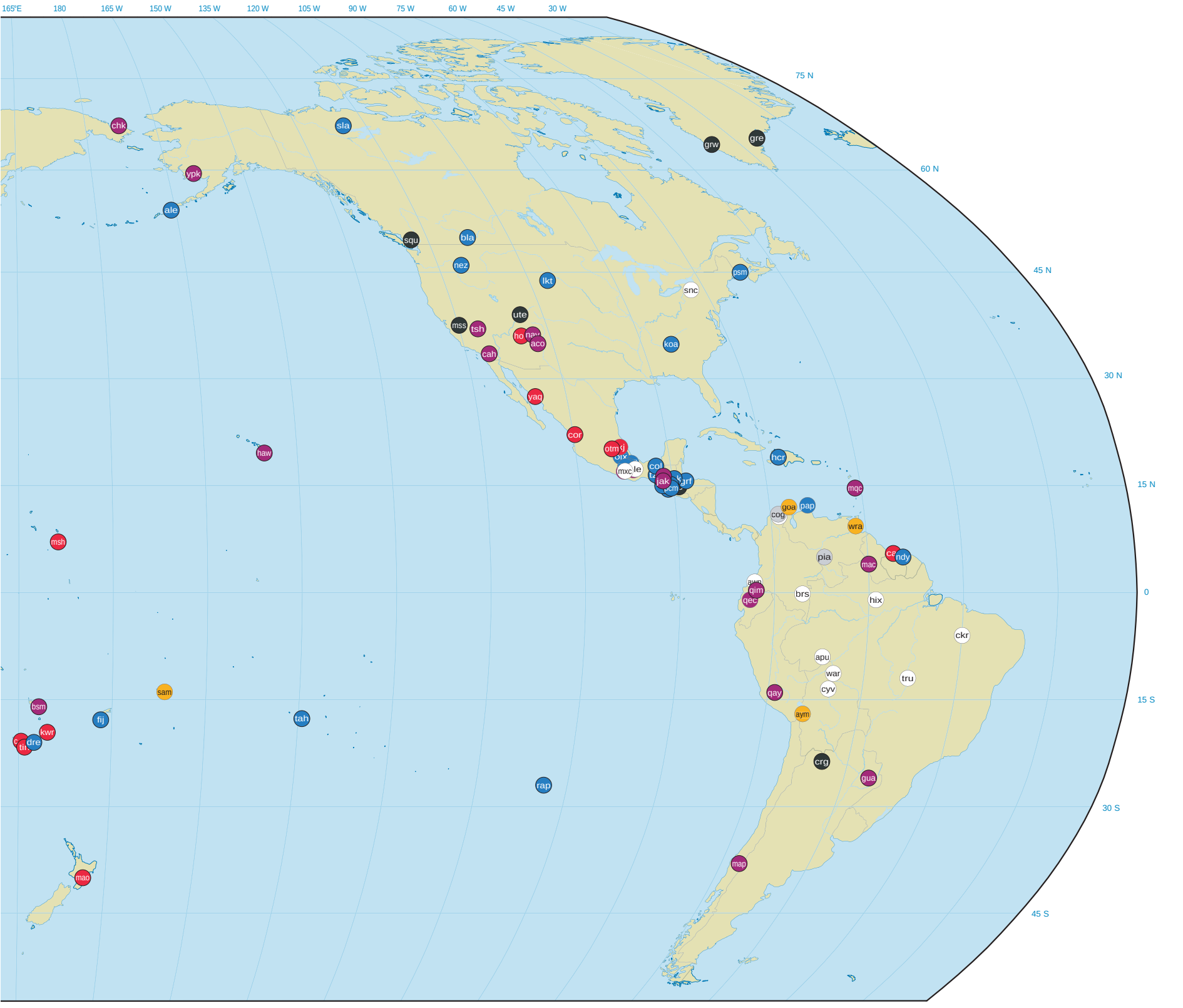


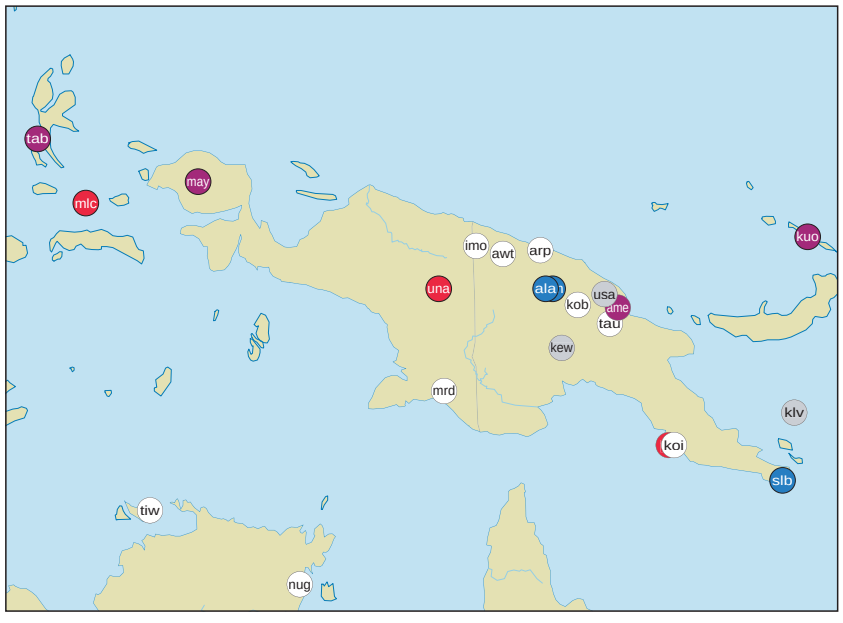


western Africa

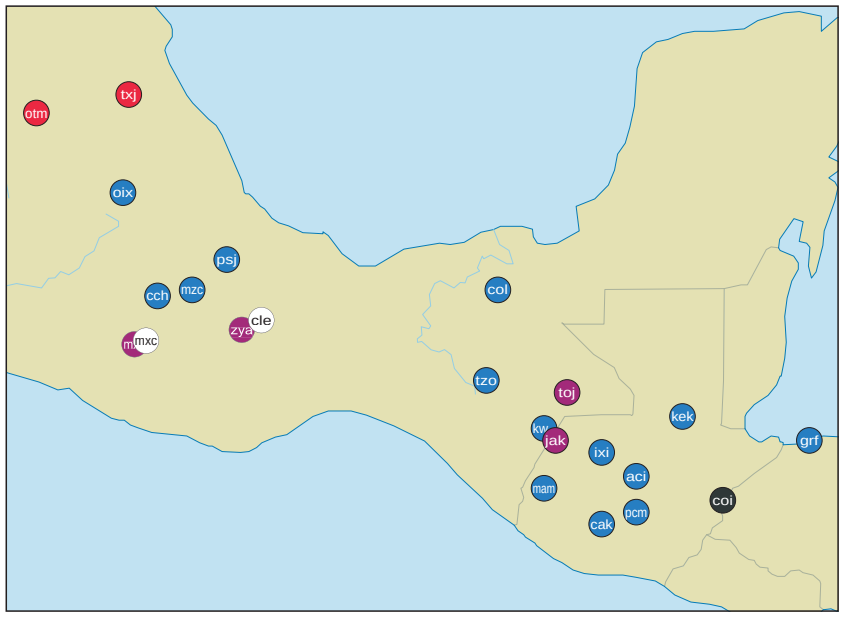


Black Sea area





New Guinea and northern Australia



Mesoamerica

54 Distributive Numerals

DAVID GIL

1 Introduction

Most languages have a series of cardinal numerals such as English *one*, *two*, *three*, and so on. Some of the ways in which such numerals are formed are described in Chapter 131. In addition, most languages also have various other series of numerals, whose forms are derived from cardinal numerals, and whose denotations combine the concept of number with other concepts of a variety of different kinds. Perhaps the most well known of these other series are the ordinal numerals, described in Chapter 53. This map is concerned with one of the many other kinds of numeral series, namely, **distributive numerals**.

Consider the following English sentence, containing the cardinal numeral *three*:

(1) *John and Bill carried* ***three*** *suitcases.*

How many suitcases did John and Bill carry? Sentence (1) is ambiguous, allowing for two possible answers to the preceding question. Under one interpretation, there was just a single set of three suitcases which both John and Bill carried. But under another interpretation, there were two sets of three suitcases, one set which John carried, another set which Bill carried. Or in other words, there were three suitcases per person, for a total of six suitcases.

In English, the latter interpretation can be forced by means of operators such as *each*, as in *John and Bill carried three suitcases each*, or *apiece*, as in *John and Bill carried three suitcases apiece*. However, in other languages, there is another way of increasing the availability of this interpretation, by means of a distributive numeral, as in the following example:

(2) Georgian (adapted from Gil 1988)
Romanma da Zurabma ***sam-sami***
Roman.ERG and Zurab.ERG DISTR-three.ABS
čanta *caiγo.*
suitcase.ABS carry.PST.3SG
(i) 'Roman and Zurab carried three suitcases each/apiece.'
(ii) 'Roman and Zurab carried the suitcases three by three.'

In Georgian, distributive numerals are formed from cardinal ones by reduplication. In the above example, *sam-sami* is a distributive numeral, derived from the cardinal numeral *sami* 'three'.

Typically, sentences containing distributive numerals also exhibit ambiguities, though of a different nature from those in sentences such as (1). For example, as indicated by the two different glosses for sentence (2) above, the distributive numeral *sam-sami* can mean either 'three . . . each/apiece' or 'three by three'. Under the latter interpretation, there were actually an indefinite number of suitcases —six, nine, twelve, fifteen, or perhaps even three hundred and sixty-nine—albeit a number that is divisible by three. Or, more perspicuously, there were an indefinite number of carrying activities, with three suitcases per activity.

The meanings of sentences containing distributive numerals can be described in terms of a binary semantic relationship of distributivity that obtains between an expression containing the distributive numeral, the distributive **share**, and some other expression in the sentence, the distributive **key**. This relationship may be expressed by means of the English preposition *per*, in accordance with the formula *SHARE per KEY*. For example, in the Georgian sentence (2), the distributive numeral *sam-sami* marks the direct-object NP *sam-sami čanta* as the distributive share. The range of meanings exhibited by the sentence results from different choices of distributive key. Thus, if the subject NP *Romanma da Zurabma* is key, there are three suitcases per carrier, resulting in interpretation (2i). On the other hand, if the verb *caiγo* is key, there are three suitcases per carrying, resulting in interpretation (2ii).

2 Feature values

This map shows which languages have distributive numerals, and, for languages that have them, the different ways in which distributive numerals are formed from cardinal ones.

1. No distributive numerals	62
2. Marked by reduplication	84
3. Marked by prefix	23
4. Marked by suffix	32
5. Marked by preceding word	21
6. Marked by following word	5
7. Marked by mixed or other strategies	23
total	250

In languages of the first type, there are **no distributive numerals**. One example of such a language is English; some others include Vietnamese (own fieldwork), Kalam (Andrew Pawley p. c.), and Hupda (Patience Epps p. c.). Such languages use a variety of paraphrases in order to convey the meanings that in other languages are expressed with distributive numerals.

One might wonder why the English translations of Georgian *sam-sami* in (2) are not themselves considered to be distributive numerals. In (2i), *three . . . each/apiece* is not a distributive numeral since it is not a continuous constituent. But what about the expression *three by three* in (2ii)? In fact, some sources do refer to expressions similar to these as distributive numerals. Observe, however, that *three by three* has a more limited range of interpretations than Georgian *sam-sami*; specifically, *three by three* forces the verb to be the distributive key, while ruling out the subject NP as distributive key. Most commonly, the term *distributive numeral* is reserved for numerals which allow NPs to be chosen as distributive key; this is the terminology that is followed in this chapter. Accordingly, *three by three* is not considered to be a distributive numeral, and hence, English is characterized as a language without distributive numerals.

The remaining six types contain languages that do have distributive numerals, providing a classification of their formal properties. The second type of languages are those that form distributive numerals by means of **reduplication**. The use of reduplication to express distributivity is clearly of iconic motivation: repeated copies of the numeral (for example *sam-sami*) correspond to multiple sets of objects (for example sets of three suitcases). Perhaps for this reason, reduplication emerges as the most common morphological strategy for the formation of distributive numerals across the world's languages. One example of a language in which distributive numerals are formed by reduplication is Georgian, mentioned above; other examples of such languages include Gã (Jonas Nartey p. c.), Burushaski (own fieldwork), and Comanche (Haugen forthcoming).

The next four types provide a cross-cutting classification of the distributive marker, depending on whether it is affixal (including clitics) or a separate word, and whether it precedes or follows the numeral to which it applies.

The third type contains languages in which distributive numerals are formed by **prefixation**. For example, in Tongan, distributive numerals are marked with the prefix *taki-*, as in *takitolu*, from *tolu* 'three' (own fieldwork). Other examples of such languages include Nicobarese (Braine 1970: 120), Yami (Der-Hwa Victoria Rau p. c.), and Coast Tsimshian (Boas 1911: 381).

The fourth type contains languages in which distributive numerals are formed by **suffixation**. For example, in Maricopa,

distributive numerals are marked with the suffix *-xper*, as in *xmokxper*, from *xmok* 'three' (Gil 1982). Other similar languages include Basque (Saltarelli et al. 1988: 253), Korean (Gil 1989), and Cayuvava (Key 1967: 61).

The fifth type contains languages in which distributive numerals are formed by one or more **preceding words**. For example, in German, distributive numerals are marked with the preceding word *je*, as in *je drei* from *drei* 'three'. Other examples of such languages include Russian (Gil 1993), Modern Greek (Alexis Dimitriadis p. c.), and Tuvaluan (Besnier 2000: 578–9).

The sixth type contains languages in which distributive numerals are formed by one or more **following words**. For example, in Malagasy, distributive numerals are marked with the following word *avy*, as in *telo avy* from *telo* 'three' (Fanja Nawalone Hanitry Ny Ale-Gerull p. c.). Other similar languages include Selkup (Prokof'eva 1966: 404), Ainu (Refsing 1986: 176–7), and Wai Wai (Hawkins 1998: 128).

Finally, the seventh type is a mixed bag containing languages in which distributive numerals are formed with **combinations** of the above types of strategies, or with yet **other morphosyntactic devices**. Combinations of morphosyntactic strategies may be either paradigmatic, in which case different kinds of distributive numerals are formed with different strategies, or alternatively syntagmatic, in which case a single distributive numeral is formed by two or more strategies in conjunction. The former case is instantiated by Udihe, in which distributive numerals may be formed either by reduplication or by suffixation with *-tA*; for example, *ila-ila* or *ilata* from *ila* 'three' (Nikolaeva and Tolskaya 2001: 428–9). The latter case is illustrated by Mayo, in which distributive numerals are formed by reduplication plus suffixation with the genitive *-ri*, for example *pesri-pesri* from *pes* 'two' (Foreman 1974: 86–8). The other morphosyntactic strategies include ambifixation, intercalated reduplication, and compounding. Ambifixation is instantiated by Chukchi, in which distributive numerals are formed by application of *em- -yut* (or variant thereof), as in *em-ŋire-yut* from *ŋire* 'two' (Skorik 1961: 397–9). Intercalated reduplication is illustrated by Tawala, in which distributive numerals are formed by insertion of *po* 'and' between copies of the numeral, for example *emosi po emosi* from *emosi* 'one' (Ezard 1997: 81, 156). And compounding is the strategy used by Kayardild, in which distributive numerals are formed by compounding with *wuthinda* 'lots' combined with reduplication, for example *kiyarrwuthinda kiyarrwuthinda* from *kiyarr* 'two' (Nick Evans p. c.).

3 Geographical distribution

Distributive numerals display some striking geographical patterns. To begin with, one may contrast the distribution of the first type of languages, those without distributive numerals, with the remaining types of languages, all of which have distributive numerals of various kinds. Languages without distributive numerals predominate in two parts of the world, mainland South-East Asia, and a rather curious contiguous zone which might be referred to as peripheral Europe, extending from Finnish via English, Spanish, and Modern Standard Arabic all the way to Persian. Elsewhere, languages without distributive numerals occur sporadically in New Guinea, Australia, and South America, areas which are more generally characterized by relatively underdeveloped numeral systems. Conversely, there are two very large contiguous zones in which almost all languages have distributive numerals of one kind or another. One such zone encompasses all of sub-Saharan Africa. The other covers the Eurasian landmass to the exclusion of peripheral Europe and South-East Asia; this zone also extends out, with gradually weakening force, into insular South-East Asia and the Pacific.

Among languages with distributive numerals, the particular morphosyntactic strategies that are used in their formation create a number of additional geographical patterns.

Reduplication is the predominant strategy in sub-Saharan Africa, the Caucasus and the South Asian subcontinent; moreover, it is also common in Indonesia and in North America, where it is interspersed with other strategies.

Prefixation is almost entirely limited to the Austronesian language family, where it occurs in two disparate areas, Taiwan plus the Philippines, and Fiji plus Polynesia. (It should be noted that most of the languages of mixed type in the Philippines also use prefixation, in combination with reduplication.) Outside of Austronesian, prefixation is found in the adjacent Nicobarese, and in the Pacific North-West, which, from a global perspective, is also relatively close to Austronesian.

A remarkable pattern is displayed by languages which form distributive numerals by means of a preceding word. With the single exception of Tuvaluan (in which the word in question is clearly cognate to the prefixes of other Polynesian languages), all languages of this type are found in a single compact area encompassing central and eastern Europe; a closer look at this area is provided in the first blow-up map. A particularly interesting pattern is displayed by Dutch, German, and Yiddish. Limiting attention to the standard languages, the boundary between language types passes between standard Dutch, which has no distributive numerals, and standard German, which, as mentioned above, forms distributive numerals with the preceding word *je*. However, turning to the colloquial varieties, a somewhat richer pattern emerges. In Germany and Austria, colloquial dialects differ from the standard language in that there is no distributive numeral construction; in this respect, these dialects resemble their Dutch counterparts further to the west. Only in Swiss dialects of German does the *je* construction occur colloquially. However, further to the east, dialects of German make use of other words to form distributive numerals. For example, in the Timişoara dialect of German, the quantifier *jeder* 'each' has undergone reanalysis as a distributive numeral formative (Gabriella Hermon p. c.). Similarly, in the Lodz dialect of Yiddish, the preposition *zu* 'to' has been co-opted as a marker of distributive numerals, in an obvious calque on the Slavic preposition *po* 'to' with similar function (Feygl Infeld Glezer, Wolf Krakowski p. c.). Thus, with the exception of the outlier Swiss dialects, the nonstandard varieties form a pattern similar to that of the standard languages, with absence of distributive numerals in the west giving way to distributive numerals formed by a preceding word as one moves further to the east.

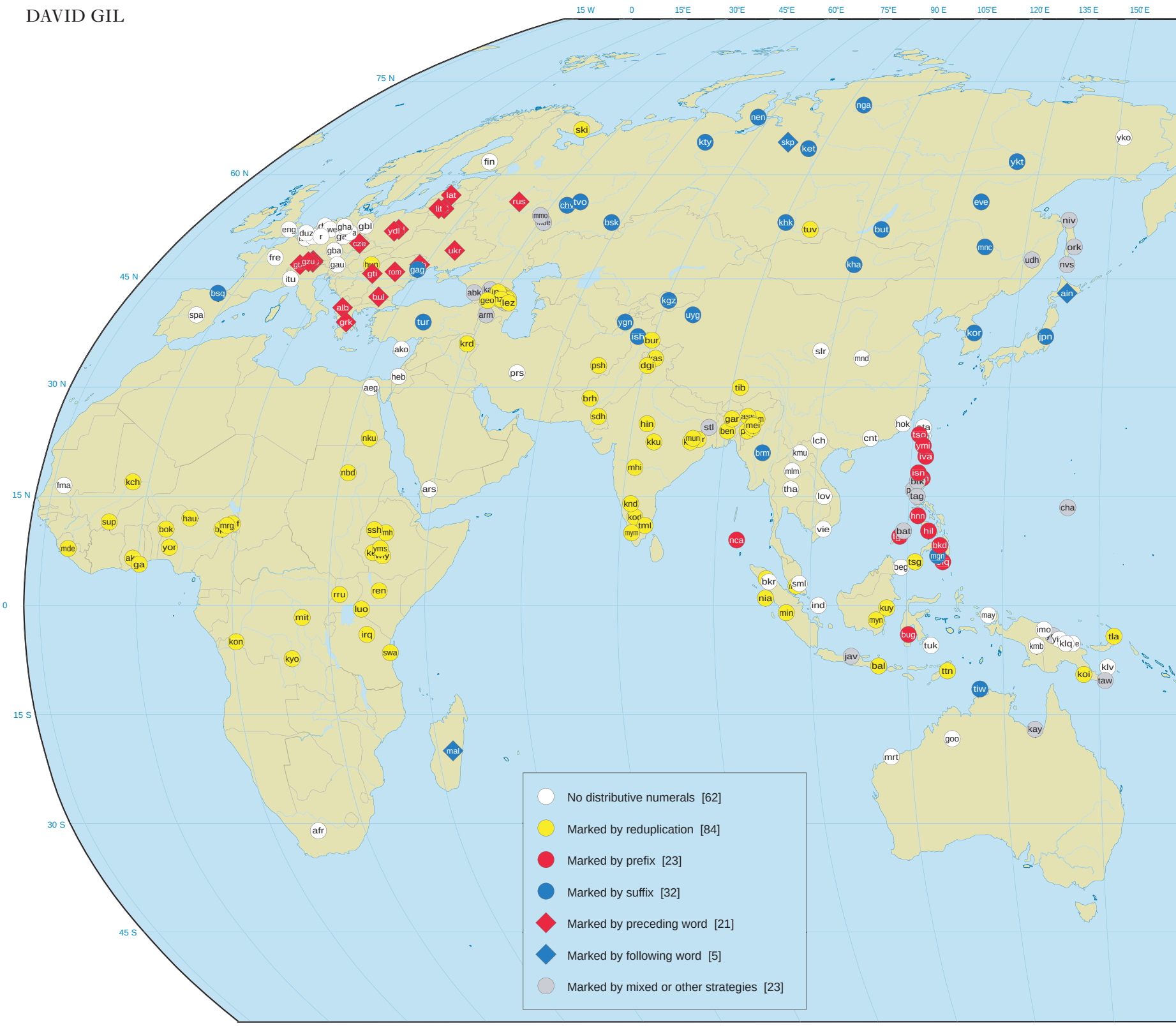
What this shows, then, is that nonstandard language varieties may exhibit geographical patterns mirroring those of their standard counterparts, albeit with a certain amount of variation. Thus, in their formation of distributive numerals by means of a preceding word, eastern dialects of German and Yiddish differ from otherwise closely related dialects to the west, instead belonging to the same isogloss as Russian and Greek. More generally, in possessing distributive numerals of one kind or another, eastern dialects of German and Yiddish lie at one end of a contiguous isogloss that stretches all the way across Asia and out into the Pacific.

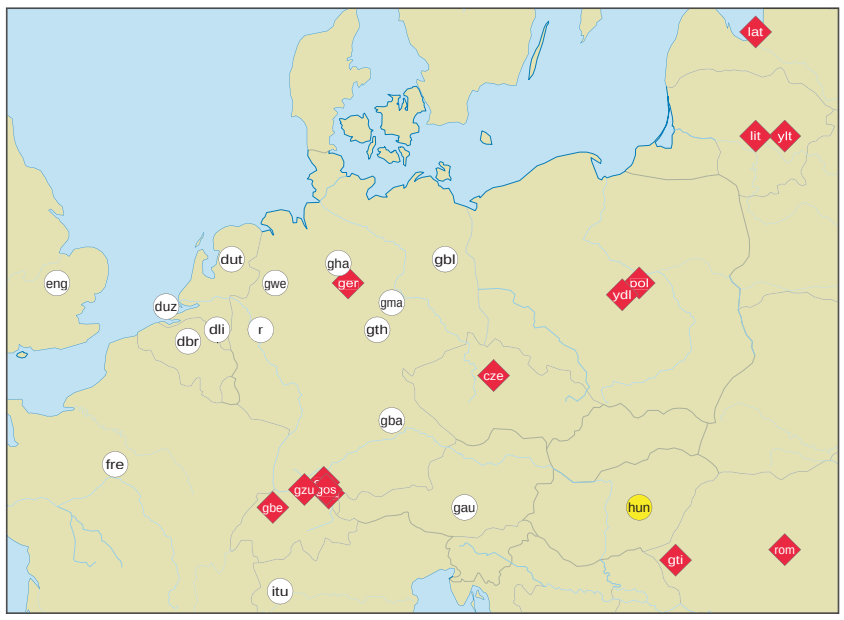
4 Theoretical issues

Distributive numerals have attracted relatively little attention in the linguistic literature. An extensive description and analysis of distributive numerals is provided in Gil (1982) and further elaborated in Gil (1987, 1988, 1989, 1992, 1993). Some attempts to account for distributive numerals within formal theories of semantics are made in Link (1986a, 1986b) and Choe (1987). The relatively minor role that distributive numerals have played within theories of natural language quantification casts doubt on the purported universality of such theories, which tend to be handicapped by the disproportionate influence of one particular language without distributive numerals spoken in the upper left-hand corner of this map, namely English.

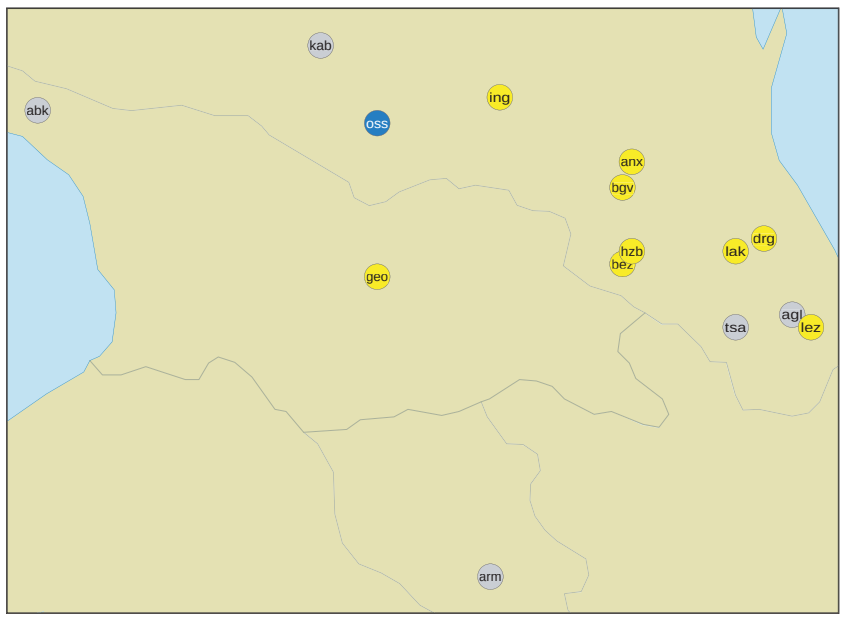
54 Distributive Numerals

DAVID GIL

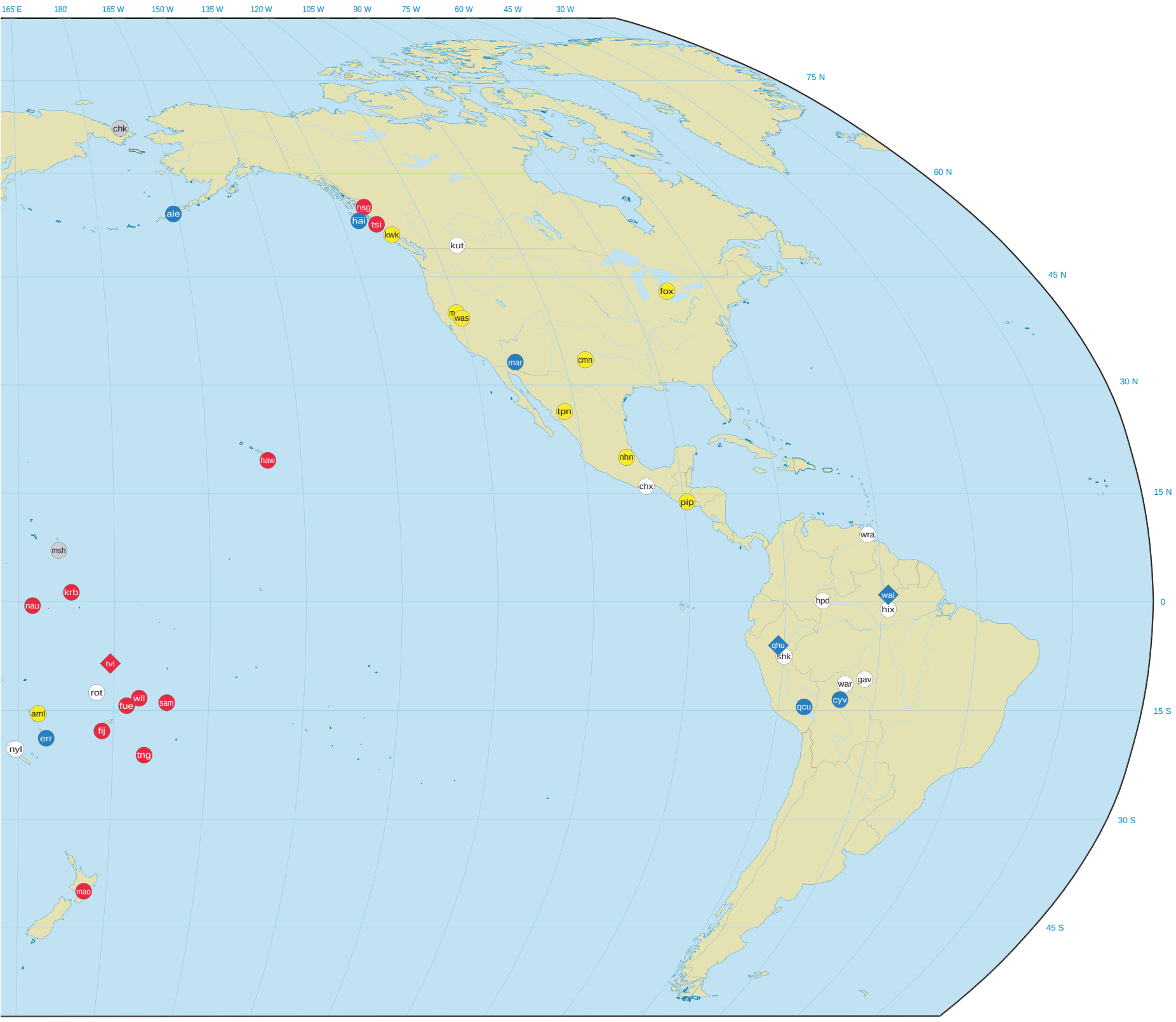


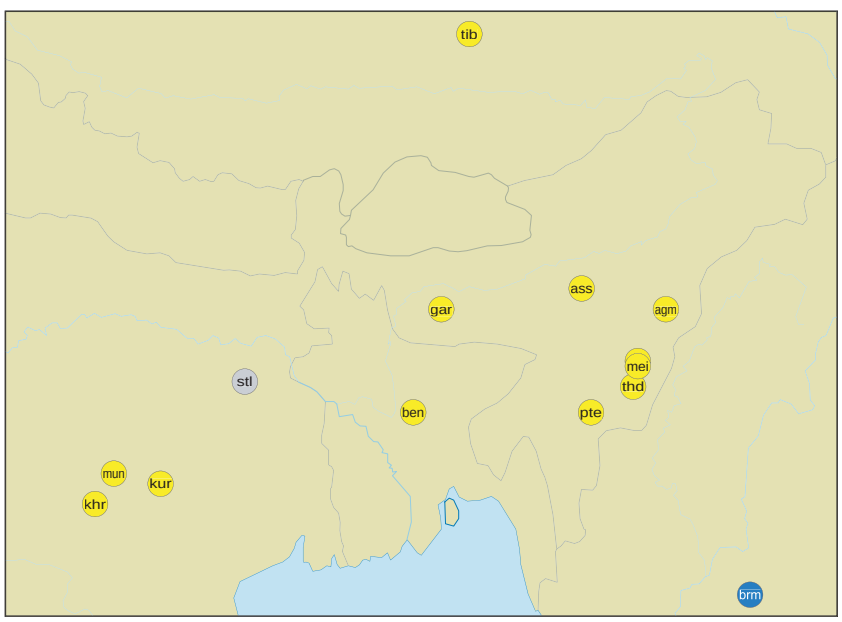


central Europe

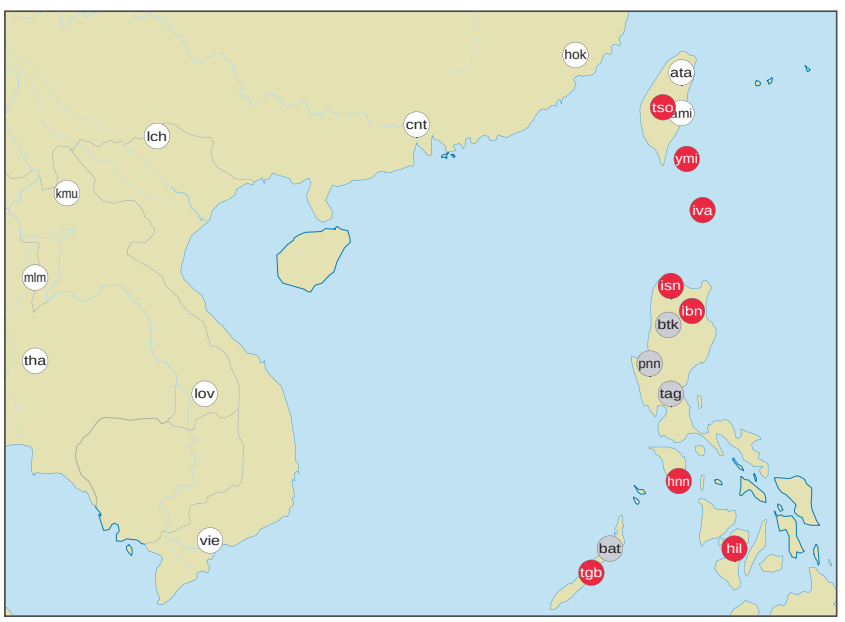


Caucasus





eastern India and surrounding areas



Indochina and Philippines

55 Numeral Classifiers

DAVID GIL

1 Introduction

In English and in other languages, nouns may vary with respect to the property of **countability**. Nouns of high countability such as *woman*, *dog*, and *pencil* denote objects which are conceptualized in terms of highly individuated units typically associated with a characteristic shape. In contrast, nouns of low countability such as *water*, *sand*, and *smoke* denote objects which are conceptualized in terms of masses without unitary structure or characteristic shape.

Countability bears a number of grammatical consequences, one of which is the way in which nouns may occur in construction with numerals. Nouns of high countability generally occur in direct construction with numerals, for example *one woman*, *two dogs*, *three pencils*. In contrast, nouns of low countability typically do not occur in direct construction with numerals; instead, an additional item must be present, for example *one **glass** of water*, *two **pounds** of sand*, *three **plumes** of smoke*. Such additional items are sometimes referred to as **mensural numeral classifiers**, since they provide nouns of low countability with a unit of measure by means of which they may then be counted.

However, in many other languages, nouns in construction with numerals may occur with an additional grammatical element even when such nouns are of high countability:

(1) Minangkabau (Sundic, Austronesian; Indonesia; own knowledge)

a. *sa-**urang*** *padusi*
one-CLF woman
'one woman'

b. *duo* ***ikue*** *anjiang*
two CLF dog
'two dogs'

c. *tigo* ***batang*** *pituluik*
three CLF pencil
'three pencils'

Such elements are typically referred to as **sortal numeral classifiers**. This term makes reference to one of the most salient functions of such forms, which is to divide the inventory of count nouns into semantic classes, each of which is associated with a different classifier. Thus, for example, in Minangkabau, as illustrated in (1) above, human nouns take the classifier *urang*, nonhuman animate nouns the classifier *ikue*, and elongated objects the classifier *batang*. In languages with sortal classifiers, the number of such classifiers may vary considerably, ranging from a mere handful to several dozen.

This map displays the worldwide distribution of sortal numeral classifiers. The map is not concerned with mensural numeral classifiers, which most or all languages have. Accordingly, sortal numeral classifiers are referred to here as *numeral classifiers* or simply as *classifiers*, in contexts where the meaning is clear.

2 Feature values

The map divides the world's languages up into three types, in accordance with whether, in the numeral-plus-noun construction, numeral classifiers are absent, optional, or obligatory.

1. Numeral classifiers are absent	260
2. Numeral classifiers are optional	62
3. Numeral classifiers are obligatory	78
total	400

In languages of the first type, there are no numeral classifiers; a numeral always occurs in direct construction with a noun without the additional presence of a classifier. One example of such a language is English. Another is Tagalog, in which the numeral-plus-noun construction is marked with the further presence of a ligature of morphophonemically variable form, in the following example the enclitic *=ng*:

(2) Tagalog (own knowledge)
dalawa=ng *aso*
two=LIG dog
'two dogs'

Other examples of languages without numeral classifiers include Yoruba (Welmers 1973: 289), Hebrew (own knowledge), Tiwi (Osborne 1974), Maricopa (own fieldwork), and Shipibo-Konibo (Pilar Valenzuela p. c.).

In languages of the second type, numeral classifiers are present, but their occurrence in the numeral-plus-noun construction is optional. One example of such a language is Minangkabau, in which classifiers such as those illustrated in (1) are sometimes present, but in other cases absent:

(3) Minangkabau (own knowledge)
duo *(**ikue**)* *anjiang*
two (CLF) dog
'two dogs'

Other examples of languages with optional numeral classifiers include Hungarian (Beckwith 1992), Chantyal (Noonan 2000), Hatam (Reesink 1999: 57–8), Tongan (own fieldwork), and Haida (Swanton 1911: 277–8).

In languages of the third type, numeral classifiers are obligatory in the numeral-plus-noun construction: a numeral cannot quantify a noun without such a classifier being present. One example of such a language is Vietnamese:

(4) Vietnamese (own fieldwork)
hai ***con*** *chó*
two CLF dog
'two dogs'

Other examples of languages with obligatory numeral classifiers include Boko (Jones 1998: 128), Nivkh (Daniel Abondolo p. c.), Nyelâyu (Françoise Ozanne-Rivierre, Isabelle Bril p. c.), Coast Tsimshian (Boas 1911: 396–8), and Warekena (Aikhenvald 1998: 298).

It is sometimes the case that in languages with numeral classifiers, their occurrence in the numeral-plus-noun construction is dependent on the choice of the numeral or of the noun. For example, with regard to numerals, in some languages numeral bases such as 'ten', 'hundred', 'thousand', and the like are themselves treated as numeral classifiers, and therefore take the place of another numeral classifier which might otherwise have been present. For the purposes of this map, classifiers are treated as obligatory if there exists a major class of numerals, such as for example two through nine, for which the numeral classifier is obligatory, even if for other numerals the classifier may be optional or absent. Similarly, in the absence of a major class of numerals for which a numeral classifier is obligatory, classifiers are treated as optional if there exists a major class of numerals for which the numeral classifier is optional, even if for other numerals it may be absent. With regard to nouns, in some languages different noun classes may differ with respect to the requirement of a numeral classifier in the numeral-plus-noun construction. For example, in Toqabaqita (Oceanic, Austronesian; Solomon Islands) only a small and semantically heterogeneous set of nouns may occur

with classifiers; most other nouns do not (Frank Lichtenberk p. c.). For the purposes of this map, classifiers are treated as obligatory if for all major classes of nouns, the classifier is obligatory. Alternatively, if this condition is not met, classifiers are treated as optional if there exists at least one class of nouns for which they are either optional or obligatory, as is the case for Toqabaqita.

The distinction between optional and obligatory numeral classifiers is further complicated by the existence of specific contexts and special styles of speech in which the distribution of numeral classifiers diverges from the norm. For example, grammar books of Vietnamese usually claim that numeral classifiers are obligatory in the numeral-plus-noun environment, and speakers of Vietnamese will typically judge the absence of a numeral classifier in such constructions to be ungrammatical. Accordingly, as noted above, Vietnamese is characterized as a language of the third type, with obligatory numeral classifiers. Nevertheless, there is at least one common context in which numeral classifiers are frequently absent, that associated with food stalls and restaurants. In a simple eating establishment, there might be a choice between two kinds of noodle soup, containing either chicken or beef. A group of, say, five people placing their order might ask for *ba gà hai bò*, or 'three chicken two beef', without any classifiers; and the person taking the order will shout back *ba gà hai bò* to the cook, again without any classifiers. For the purposes of the present map, such specific contexts and specialized styles of speech are not taken into consideration; the usage of numeral classifiers is described as it occurs in what might be characterized as "ordinary citation" language.

3 Geographical distribution

Numeral classifiers exhibit striking worldwide distribution at the global level. The main concentration of numeral classifiers is in a single zone centred in East and South-East Asia, but reaching out both westwards and eastwards. To the west, numeral classifiers peter out as one proceeds across the South Asian subcontinent; thus, in this particular region, the occurrence of numeral classifiers cross-cuts what has otherwise been characterized as one of the classical examples of a linguistic area, namely, South Asia. However, numeral classifiers pick up again, albeit in optional usage, in parts of western Asia centring on Iran and Turkey; it is not clear whether this should be considered as a continuation of the same large though interrupted isogloss, or as a separate one. To the east, numeral classifiers extend out through the Indonesian archipelago, and then into the Pacific in a grand arc through Micronesia and then down to the south-east, tapering out in New Caledonia and western Polynesia. Interestingly, whereas in the western parts of the Indonesian archipelago numeral classifiers are often optional, in the eastern parts of the archipelago and in Micronesia numeral classifiers tend once more, as in mainland East and South-East Asia, to be obligatory.

Outside this single large zone, numeral classifiers are almost exclusively restricted to a number of smaller hotbeds, in West Africa, the Pacific North-West, Mesoamerica, and the Amazon basin. In large parts of the world, numeral classifiers are completely absent.

4 Theoretical Issues

Numeral classifiers have captured the imagination of linguists and others interested in language for a variety of reasons. Most obviously, linguists, psychologists, anthropologists, and other scholars have been interested in the classificatory function of numeral classifiers and what insights it might provide into how speakers view the world (Haas 1942; Burling 1965; Craig 1986). Whereas in some languages each noun is associated with exactly one classifier, in other languages there is a degree of flexibility with regard to the choice of classifier, permitting the differentiation of subtle shades of meaning. For example, in Mandarin, when *sān* 'three' occurs in construction with *huā* 'flower', *sān* ***zhī*** *huā*, with classifier *zhī* for elongated objects, conjures up an image of three flowers on their stalks, whereas *sān* ***duǒ*** *huā*, with classifier *duǒ* for round objects, focuses in on the actual florescences, to the exclusion of the stalks.

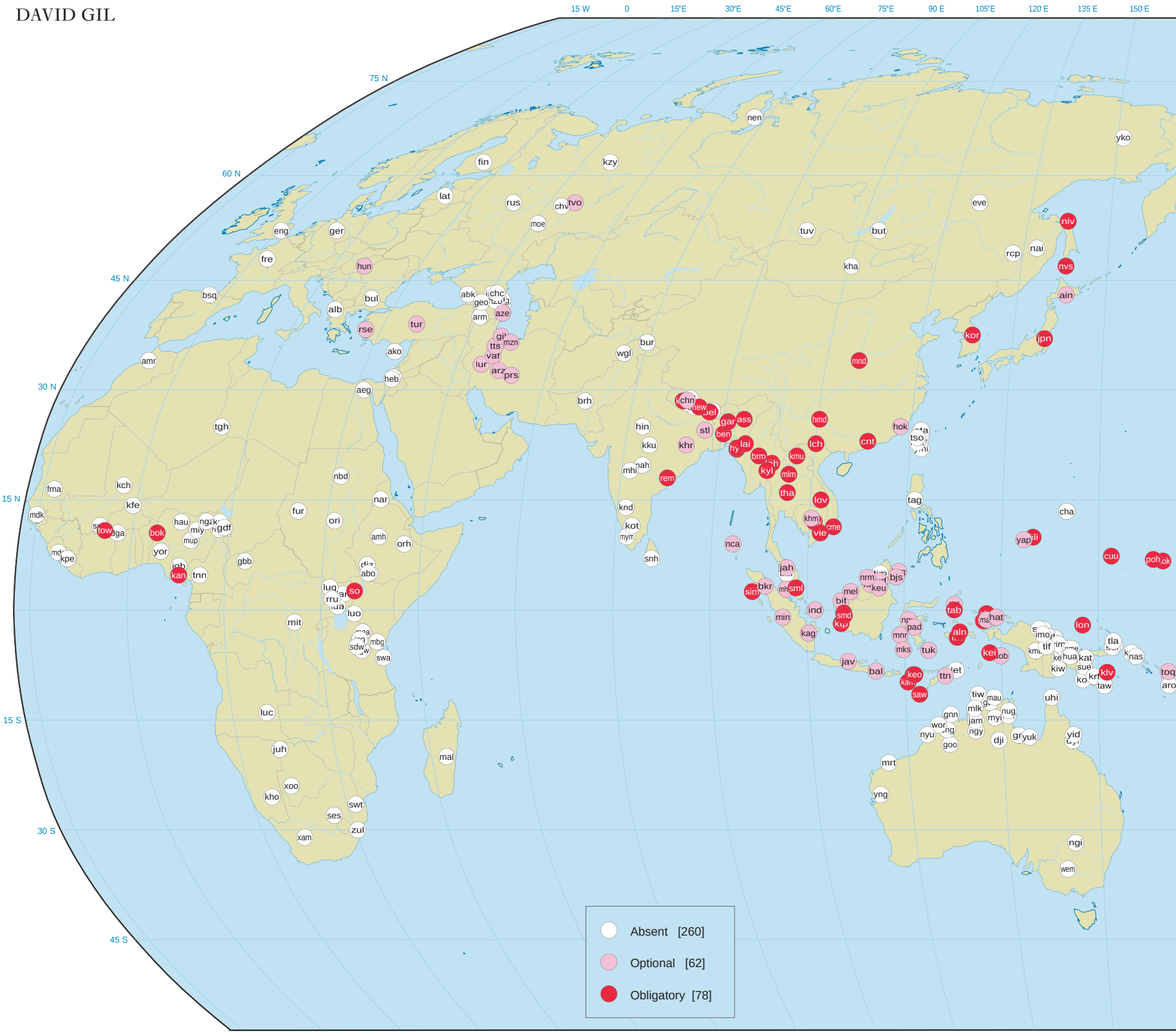
Other interesting issues pertain to the grammar of numeral classifiers. One question is that of constituency: does the classifier form a constituent with the numeral or with the noun? There is at least some reason to believe that the answer may vary across languages, and also across different kinds of classifiers. A second question is that of headedness: given a particular constituency, which of the elements is the head and which is the modifier? Again, different scholars have proposed different answers to this question, which may be due to differences between languages, differences between theoretical approaches, or perhaps other reasons.

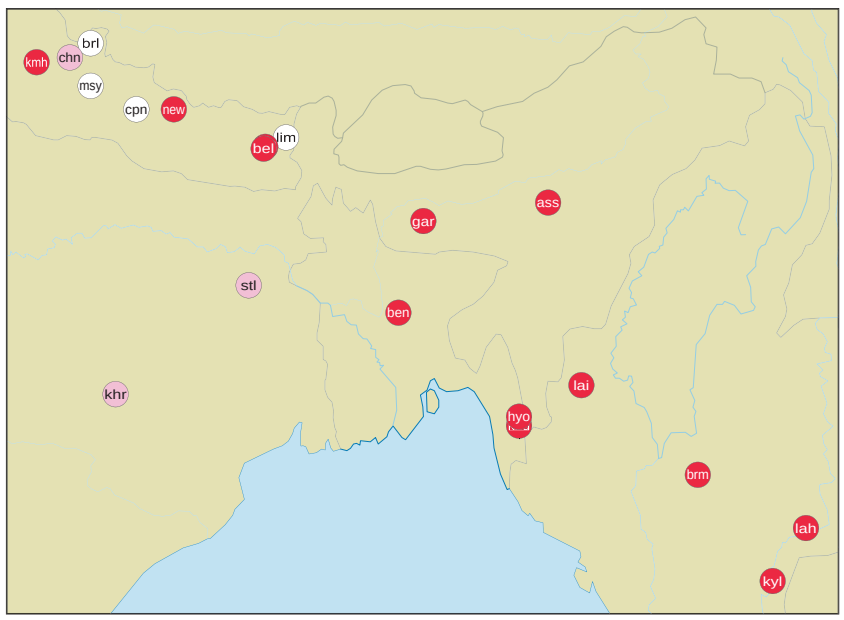
Whereas in some languages, for example Minangkabau, numeral classifiers occur only with numerals, in other languages they may occur in a variety of other constructions. For example, in Mandarin they may also occur with several other quantifiers as well as with demonstratives, as in *zhèi ge píngguǒ* 'this CLF apple', while in Cantonese, in addition to the above, they may occur also with possessor noun phrases, as in a^{33}*-faay*^{55} ***laʔ***^{5} *piŋ*^{11}*-guo*^{35} 'Ah Fai's apple'. In addition, numeral classifiers sometimes have a nominalizing function of the kind described in Chapter 61. For example, in Vietnamese, they may be used to nominalize demonstratives, adjectives, and relative clauses, as in ***con*** *Quân dã mua* 'one that Quan bought'. Finally, in some languages, numeral classifiers may occur in construction with a noun without any other attributive expression being present; in such cases, the classifier functions rather like an article, though varying from language to language with respect to definiteness.

Why do languages have numeral classifiers? Most commonly, classifier languages are ones in which an NP consisting entirely of a bare noun may be understood as either mass or count, and as either singular or plural (Gil 1987). For example, in Mandarin, the noun phrase *píngguǒ* 'apple' may be understood as 'apple' (mass), 'an apple' (singular), or 'some apples' (plural). This cross-linguistic generalization has led to a widespread belief that in such languages, the classifier is needed in order to individuate the noun and provide the necessary units to facilitate quantification (Thompson 1965; Quine 1969; Stein 1981; Link 1991). Thus, according to this view, the Mandarin **sān píngguǒ* 'three apple(s)' is semantically ill formed for the same reason that the English **three water(s)* is: just as English *water* requires an explicit mensural classifier before it can be quantified, as in *three glasses/ounces/drops of water*, so Mandarin *píngguǒ* requires a sortal classifier before it can be successfully enumerated, as in *sān gè píngguǒ*. However, this explanation for the occurrence of classifiers is belied by a variety of additional facts (Gil 1996). Unlike English *water*, the Mandarin noun *píngguǒ* does indeed come with a conventional unit of enumeration, namely the individual round fruit that one can hold in one's hand, and therefore does not need a classifier in order to be counted. Evidence for this is provided by size and shape adjectives: whereas in English, constructions such as *big water* are bizarre, in Mandarin, constructions such as *dà píngguǒ* 'big apple' are syntactically well formed, and understood in the same way as their English counterparts. A further argument against this explanation is provided by the existence of languages in which, as in Mandarin and other classifier languages, an NP consisting entirely of a bare noun may be understood as either mass or count, and as either singular or plural, but in which there are no numeral classifiers. For example, in Tagalog, the noun phrase *mansana* 'apple' may be understood as 'apple' (mass), 'an apple' (singular), or 'some apples' (plural); however, it does not require a classifier in order to occur with a numeral, as in *tatlong mansana* 'three apples'. Thus, the occurrence of classifiers in classifier languages has no straightforward semantic explanation; in the absence of any better-motivated account, it must accordingly be viewed as just another arbitrary, conventionalized fact about the grammars of individual languages.

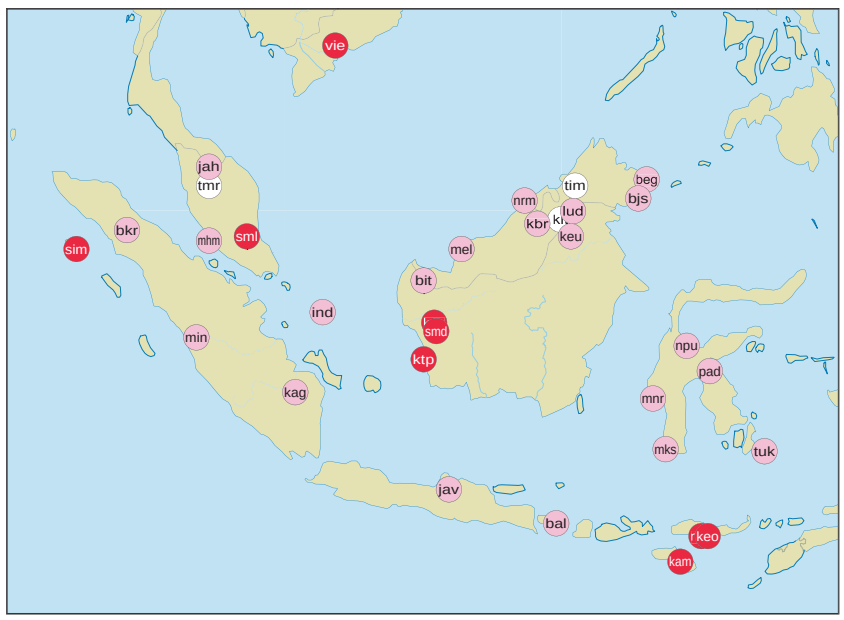
55 Numeral Classifiers

DAVID GIL

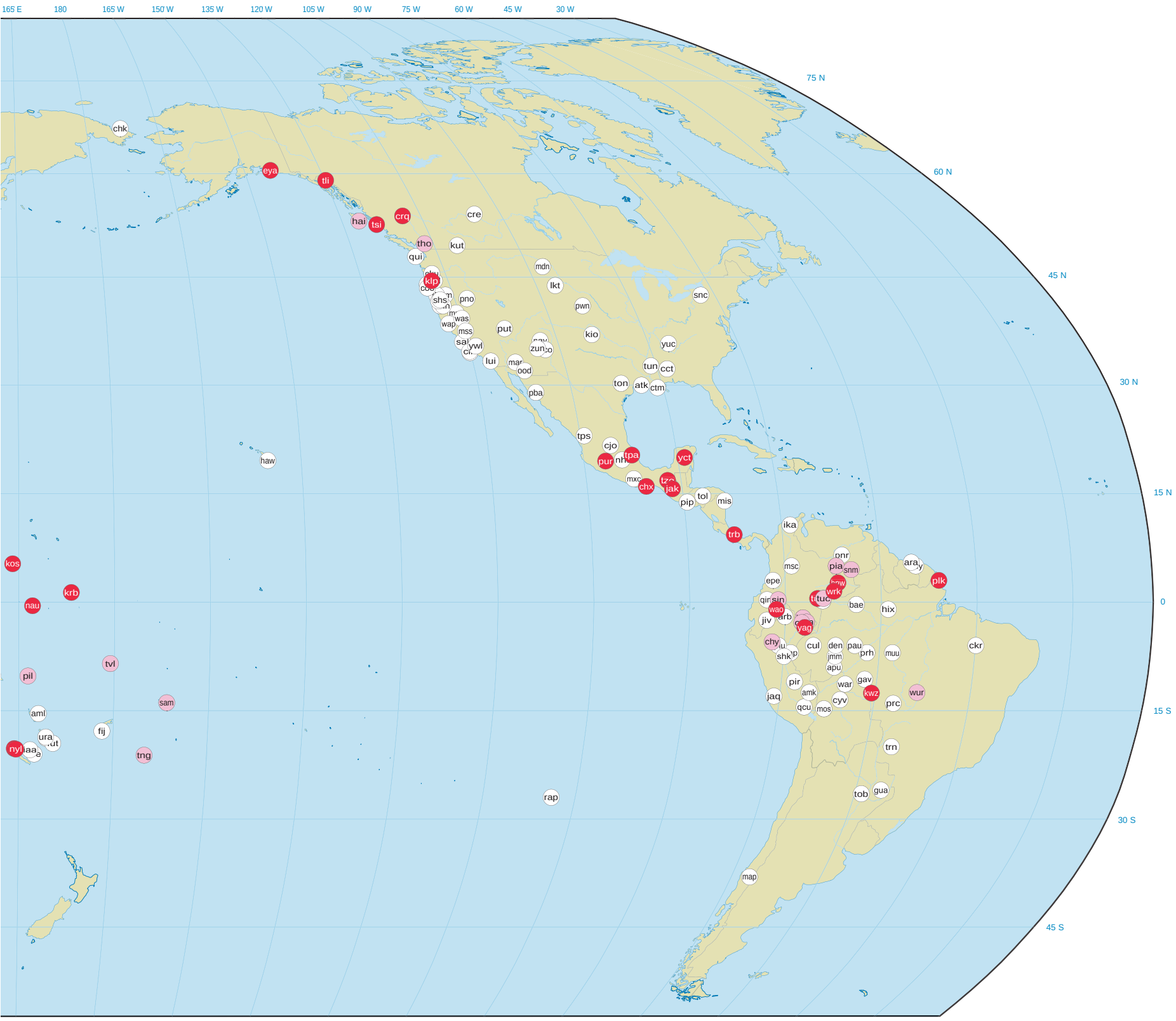


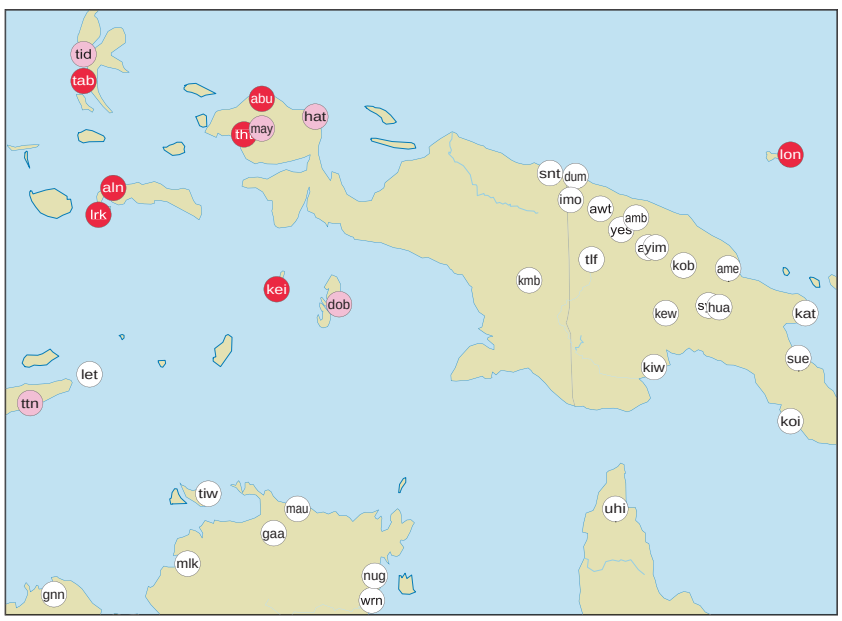


Nepal, eastern India, and surrounding areas

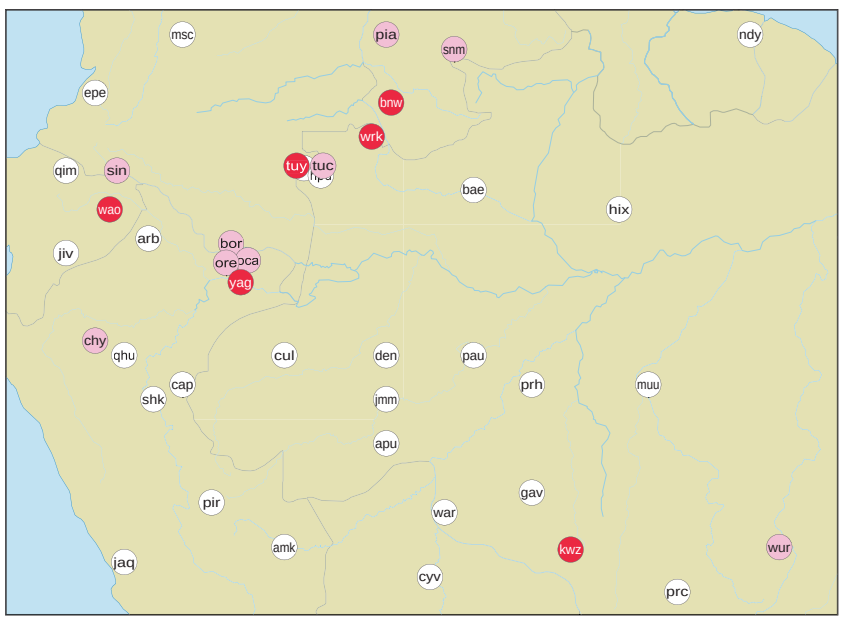


western Indonesia





New Guinea and northern Australia



northwestern South America

56 Conjunctions and Universal Quantifiers

DAVID GIL

1 Introduction

Conjunctions are forms with meanings similar to that of English *and*. **Universal quantifiers** are expressions with meanings resembling those of English *every*, *each*, and *all*.

Logicians, as well as logically minded linguists, have suggested that there is a close affinity between conjunctions and universal quantifiers. For example, in the context of a class consisting of five students, Alice, Bill, John, Mary, and Susan, sentence (1) with the conjoined NP *Alice, Bill, John, Mary, and Susan* is logically equivalent to sentence (2) with the universally quantified noun phrase *every student*.

(1) *Alice, Bill, John, Mary,* ***and*** *Susan passed the exam.*
(2) ***Every*** *student passed the exam.*

Based on observations such as these, some semanticists have proposed deriving the interpretations of universal quantifiers from those of conjunctions. For example, in the Boolean Semantics of Keenan and Faltz (1986), conjunctions and universal quantifiers are both represented in terms of set-theoretic intersections.

How well do such semantic representations correspond to the observable lexical and grammatical patterns of languages? On the basis of examples such as (1) and (2) above, one might suspect that they do not correspond at all well. Thus, in English, the conjunction *and* and the universal quantifier *every* are distinct words with quite different grammatical properties.

However, a broader cross-linguistic perspective suggests that there are indeed widespread lexical and grammatical resemblances between conjunctions and universal quantifiers, thereby lending support to the logicians' analyses. The purpose of this map is to portray some of these connections, and, in doing so, to show how the cross-linguistic study of such lexical and grammatical patterns can be of relevance to logicians and their theories of formal semantics.

2 Feature values

For the purposes of the map, conjunctions are taken to include not only forms with meanings similar to that of *and*, but in addition expressions that are sometimes characterized as **conjunctive operators** or **focus particles**, with meanings resembling those of *also*, *even*, *another*, *again*, and in addition the restrictive *only*. As for universal quantifiers, these are assumed to encompass not only forms with meanings such as those of *every*, *each*, and *all*, but also expressions that are sometimes referred to as **free choice**, with meanings corresponding to that of *any* in constructions such as *Any student can pass the exam* (but not constructions such as *Alice didn't see any students*, where *any* has a so-called negative polarity interpretation).

1. Formally different	40
2. Formally similar, not involving interrogative expression	33
3. Formally similar, involving interrogative expression	43
total	116

The map distinguishes between three types of languages. The first type contains languages in which there is no formal resemblance between any of the conjunctions and any of the universal quantifiers. The second type contains languages in which such resemblances, which may be of variegated kinds, do exist. The third type, a specific subtype of the second type involving a specific kind of resemblance, contains languages in which universal quantifiers are formed from a combination of conjunctions and interrogative expressions.

The first type of language is exemplified by French. In French, the inventory of conjunctions consists of *et* 'and', *aussi* 'also', *même* 'even', *autre* 'other', *encore* 'again', *seulement* 'only', and others. And the inventory of universal quantifiers consists of *tout* 'all', *chaque* 'every', *n'importe quel* 'any', and others. There are thus no observable resemblances between these two classes of words. Other languages belonging to this type include Lango, Brahui, Jaminjung, Kutenai, and Panare.

The second type of language is of a heterogeneous nature, due to the many different ways in which conjunctions and universal quantifiers may be formally related. The most obvious way is through complete identity. For example, in Supyire, the form *mú* has a range of meanings that includes the conjunctive 'also' and the universal quantifier 'all' (Carlson 1994: 686). In Yidiny, the suffix *-bi* has a range of meanings that includes the conjunctive 'another' and the universal quantifier 'all' (Dixon 1977: 147–8). And in Coast Tsimshian, the prefix *max-* has a range of meanings that includes the conjunctive 'only' and the universal quantifier 'all' (Boas 1911: 317).

In a larger number of cases, the formal resemblance between conjunctions and universal quantifiers is partial rather than complete. In a few cases, conjunctions and universal quantifiers contain a common root plus some additional material specific to each of the two. For example, in Malagasy, the common root *na* 'or' may combine with *aza* 'even' to yield the conjunction *na . . . aza* 'even', or with *iza* 'who' plus reduplication to yield the universal quantifier *na iza na iza* 'anybody' (Fanja Nawalone Hanitry Ny Ale-Gerull p. c.).

In a few other instances, a universal quantifier forms part of a larger conjunction. An example of this is provided by English, in which the universal quantifier *all* is at least diachronically part of the conjunction *also*.

Considerably more common, however, is the opposite state of affairs, in which a conjunction forms part of a larger universal quantifier. For example, in Iraqw, *hleemee* 'also' suffixed with feminine *-r* and "background" suffix *-o* yields the universal quantifier *hleemeero* 'all' (Maarten Mous p. c.). Similarly, in Chukchi, *əmə* 'and' plus the nominalizing suffix *-lʔo* produce the universal quantifier *əməlʔo* 'all' (Michael Dunn p. c.). And in Taba, *le* 'only' combined with the classifier *ha* and the numeral *so* 'one' results in the universal quantifier *hasole* 'all' (Bowden 2001: 183).

In a variation on the above pattern, a conjunction may combine with a simple lexical universal quantifier to create a more complex universal quantifier expression. For example, in Amele, *cunug* 'all' frequently co-occurs with *ca* 'and', 'with' (Roberts 1987). Similarly, in Haisla, *ag-* 'all' often combines with *-am* 'just', 'really' (Bach 1996). And in Hebrew, *kol* 'all', *va* 'and', plus a reduplicated head noun yield a construction of the form *kol N va-N* with the interpretation 'every N' (own knowledge).

One of the ways in which a conjunction may form part of a larger universal quantifier is of sufficient importance to merit the positing of a third type of language: in such languages, conjunctions combine with interrogative expressions to produce universal quantifiers. For example, in Kanuri, *yayé* 'even if' combines with interrogative forms such as *ndú* 'who' to produce universal quantifiers such as *ndú yayé* 'everybody' (Cyffer and Hutchison 1990: 189). Similarly, in Colloquial Singaporean English (also known as Singlish), *also* 'also' combines with interrogative expressions such as *which* 'which' to yield discontinuous universal quantifiers such as *which . . . also* 'any' (Gil 1994c). And in Jaqaru, the suffix *-psa* 'also' attaches to interrogative stems such as *kaw* 'where' to create universal quantifiers such as *kawpsa* 'anywhere' (Hardman 2000: 34–5).

In some other cases, the conjunction and the interrogative expression combine with an additional marker or markers to form the universal quantifiers. For example, in Mosetén, the suffix *-nä* 'and' attaches to the interrogative form *jäen'* 'how' plus the associative marker *-tyi'* to produce the universal quantifier *jäen'-tyi'-nä*

'anybody' (Jeanette Sakel p. c.). Often, the additional marker in question involves reduplication. For example, in Sesotho, *le-* 'and', 'with' occurs between two copies of the interrogative form *ofe* 'which' to yield the universal quantifier *ofe le-ofe* 'every' (Guma 1971). And in Begak-Idaan, *ja?* 'only', 'just' occurs after reduplicated interrogative expressions such as *nu-nu* 'what' to yield universal quantifiers such as 'any' (Nelleke Goudswaard p. c.).

Finally, it should be noted that languages of the third type overlap to a considerable degree with languages characterized as having interrogative-based indefinite pronouns in Map 46. However, the overlap between these languages is far from complete, for at least the following reasons: (i) not all interrogative-based indefinite pronouns contain a conjunction (some consist just of a bare interrogative expression); (ii) not all interrogative-based indefinite pronouns are universal quantifiers (some are existential); and (iii) not all combinations of conjunctions and interrogative expressions forming universal quantifiers are pronouns (some occur only in attributive or determiner position).

3 Geographical distribution

As is evident from the map, languages of the first two types occur all over the world, without any significant geographical patterning. Given the many different ways in which conjunctions and universal quantifiers may be formally related to each other, the absence of such patterns is hardly surprising. Nevertheless, the fact that formal resemblances between conjunctions and universal quantifiers can be found across the globe, in geographically, genealogically, and typologically unrelated languages, vindicates the logicians' analyses, providing cross-linguistic support for semantic representations which relate conjunctions and universal quantifiers.

In contrast, languages of the third type, in which conjunctions combine with interrogative expressions to form universal quantifiers, exhibit rather striking geographical patterning. Such languages can be found in a number of hotbeds throughout the world, in central western Africa, the Caucasus, and South America. More saliently, though, such languages are the rule in a large contiguous swath encompassing South, South-East, and East Asia. In Emeneau (1980) the construction in question is argued to be one of the characteristic features of the South Asian linguistic area; however, as shown in Gil (1994b) and in the present map, the isogloss is actually much larger, extending far to the east of the South Asian subcontinent. A vivid example of how languages coming into the region undergo typological adaptation is provided by the Singaporean variety of English, which, as mentioned above, has acquired the construction, presumably under the influence of Tamil, Malay, and Chinese substrates. Thus, in the following example, interrogative *which* combines with conjunctive operator *also* to form a free-choice universal quantifier meaning 'any':

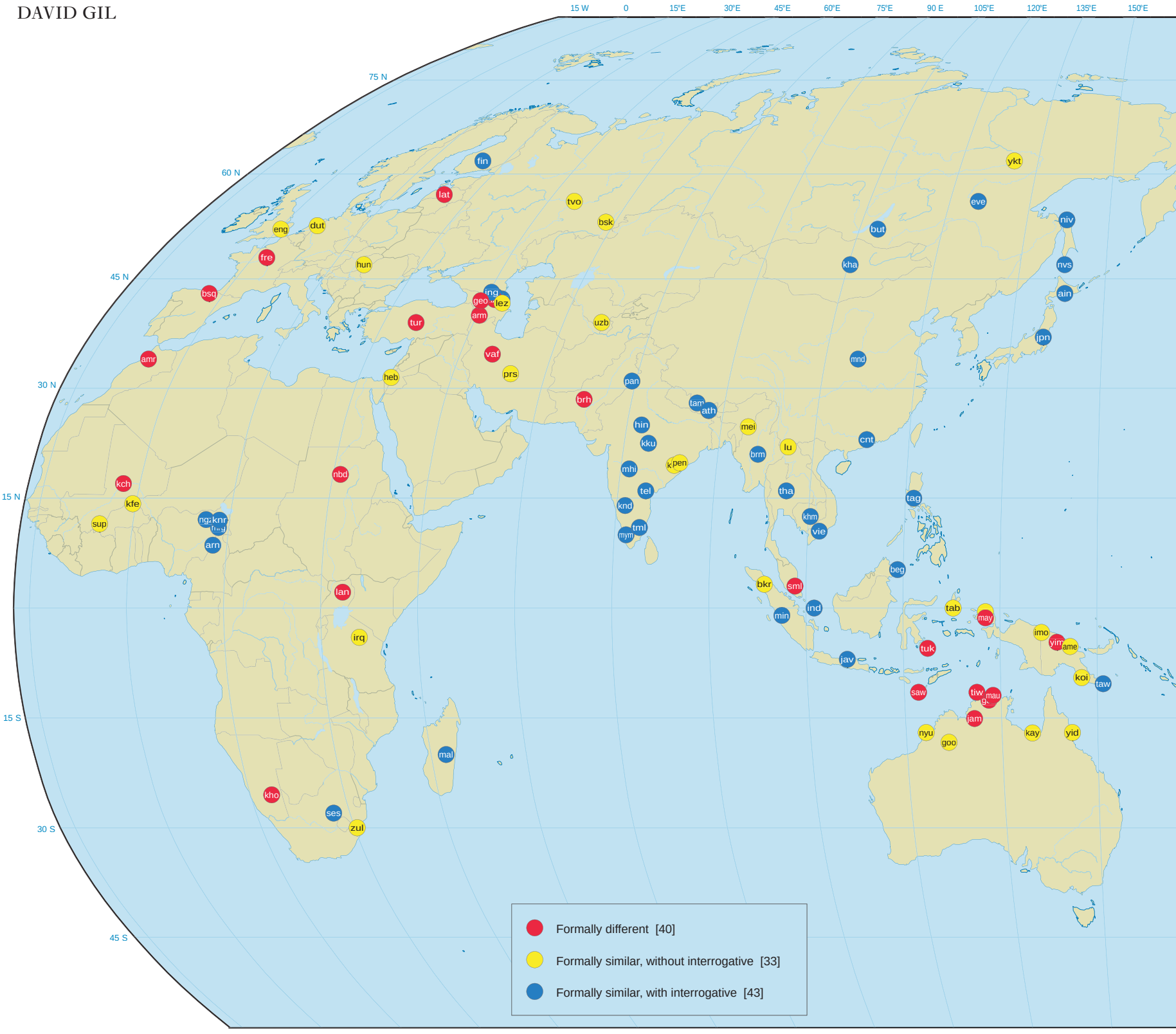
(3) Colloquial Singapore English (Singlish)
***Which** student **also** can pass the exam.*
'Any student can pass the exam.'

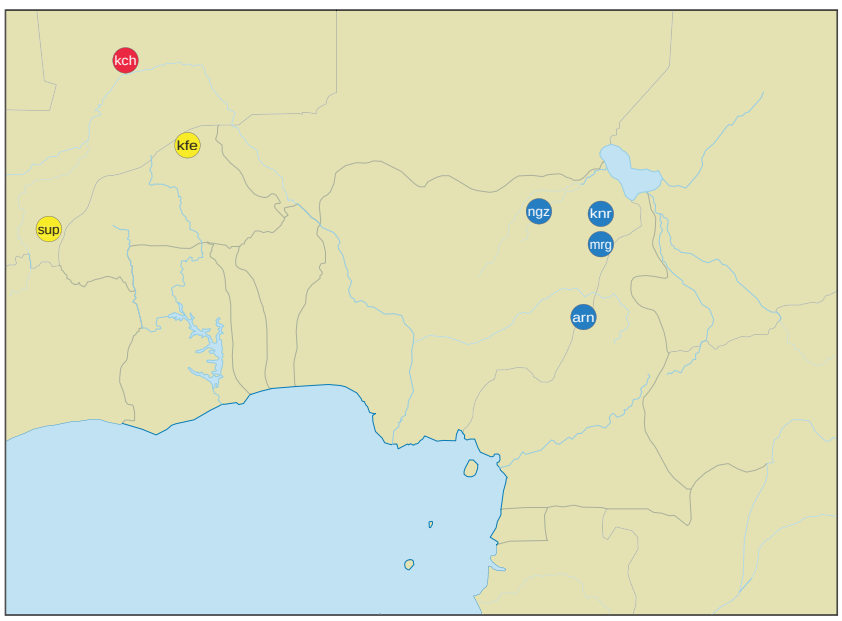
4 Theoretical issues

While the connection between conjunctions and universal quantifiers is well motivated semantically, it is still necessary to work out the detailed mechanisms by which the relevant complex expressions derive their meanings from those of their constituent parts. In particular, the construction involving the combination of a conjunction and an interrogative expression to produce a universal quantifier has been the focus of a number of recent analyses, attempting to explain how the construction acquires its resulting meaning; see, for example, König (1991), Gil (1994a,b,c), Haspelmath (1997), and others.

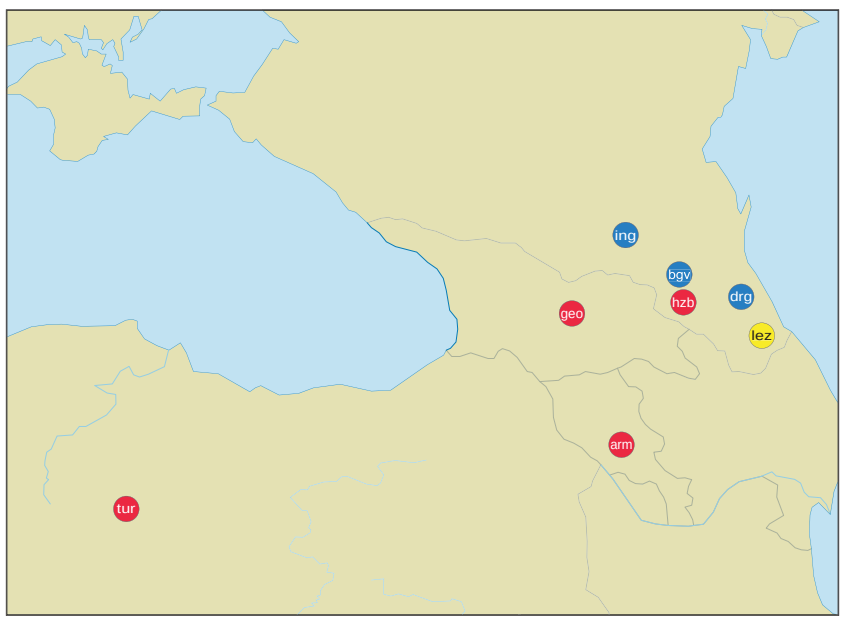
56 Conjunctions and Universal Quantifiers

DAVID GIL

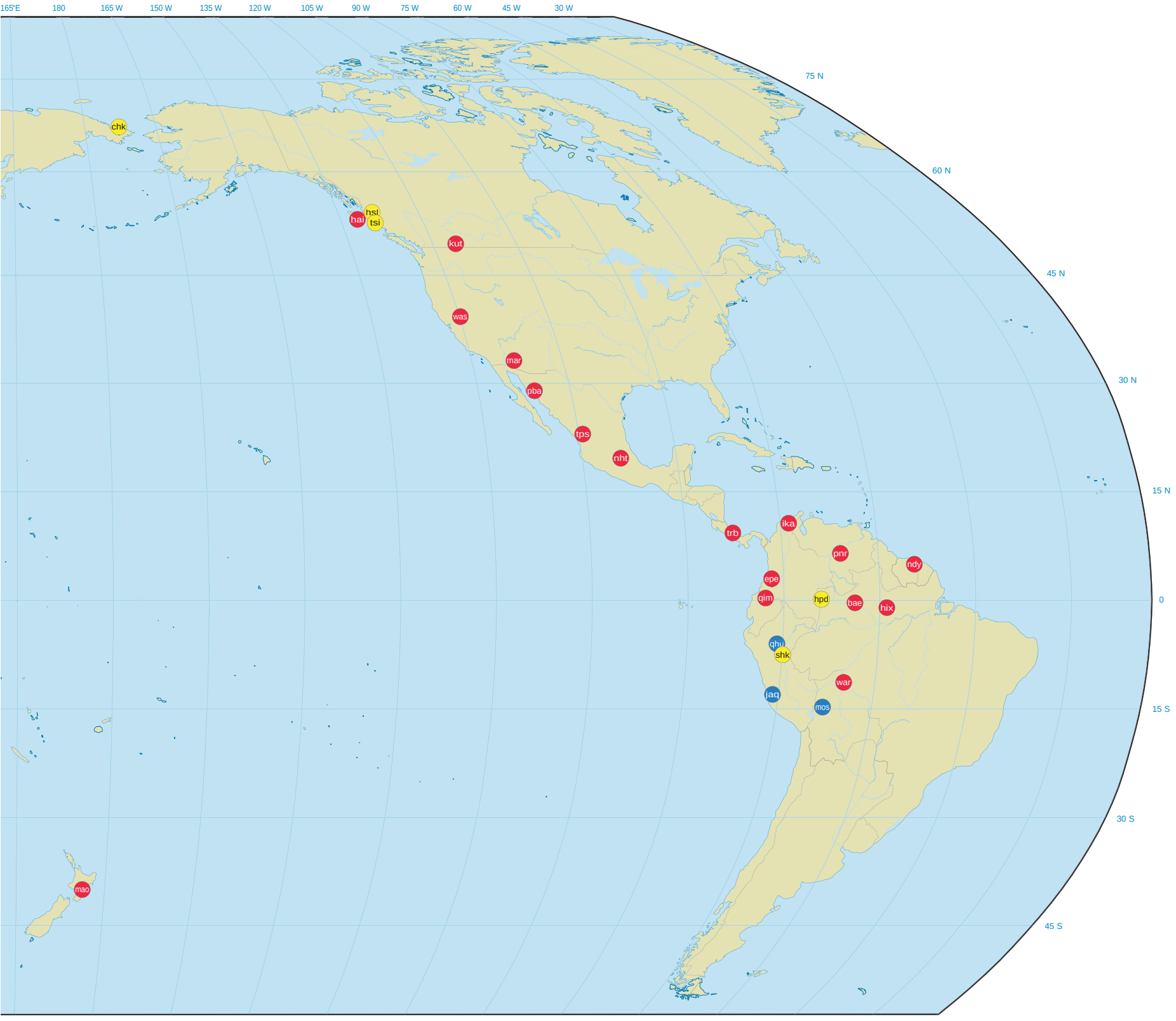


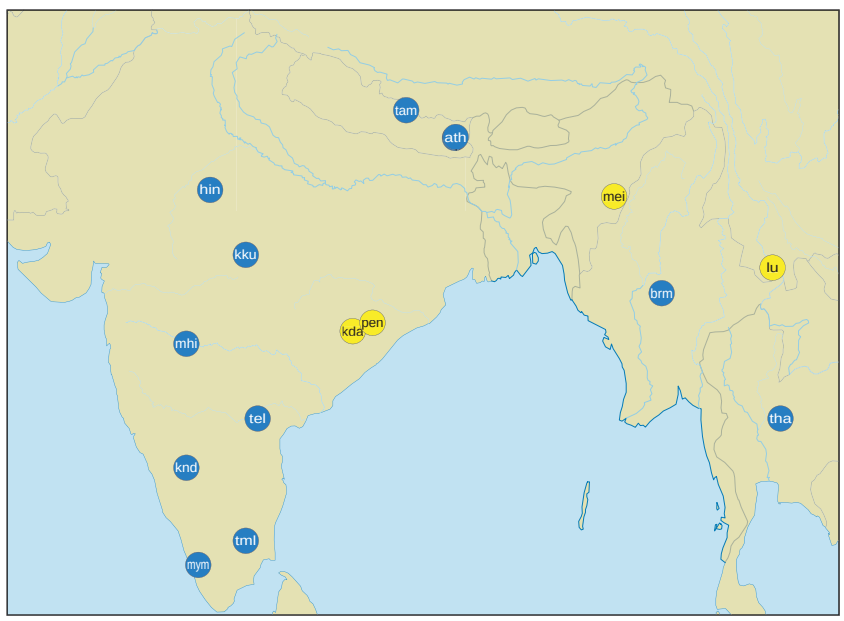


Nigeria and surrounding areas

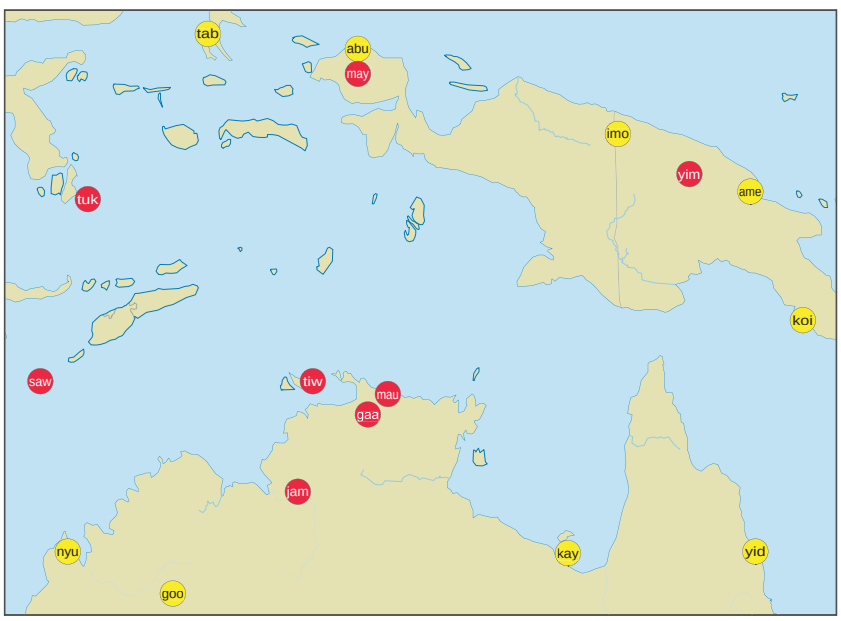


Caucasus and adjacent areas





from India to Thailand



New Guinea and northern Australia

57 Position of Pronominal Possessive Affixes

MATTHEW S. DRYER

1 Defining the values

This map shows the position of possessive affixes on nouns, as illustrated by the first-person possessive prefix *u-* in (1a) from Macushi (Carib; Guyana etc.) and by the second-person singular possessive suffix *-m* in (1b) from Loniu (Oceanic, Austronesian; Manus Island, Papua New Guinea).

(1) a. Macushi (Abbott 1991: 86)
u-pana'
1.POSS-ear
'my ear'

b. Loniu (Hamel 1994: 46)
ɛyɛ-m
blood-2SG.POSS
'your blood'

1.	Possessive prefixes	238
2.	Possessive suffixes	330
3.	Both possessive prefixes and possessive suffixes, with neither primary	24
4.	No possessive affixes	203
	total	795

Note that the term **possessive affix** is used by different linguists for two very different sorts of affixes. This chapter uses the term for affixes like those in (1) which code the person, number, and/or gender of a possessor, and which appear on a noun denoting the entity which is possessed. The term is occasionally used for what will be called here **genitive affixes**, case affixes which occur on nouns denoting possessors, to indicate that these nouns denote the possessor of something expressed by a separate noun, as illustrated by the example in (2) from Archi (Nakh-Daghestanian; Russia).

(2) Archi (Kibrik 1994: 312)
diya-n *k'oč'o*
father-GEN cup
'father's cup'

Genitive affixes have the same function as the preposition *of* or the clitic *'s* (as in *John's car*) in English; possessive affixes have the same meaning as English words like *my, your*, and *their*. The contrast between possessive affixes and genitive affixes is brought out clearly by languages like Québec-Labrador Inuktitut (Eskimo-Aleut), which employ both in genitive constructions. In (3), there is both a genitive case affix on the noun *illu* 'house' and a third-person possessive affix on the noun *ukkua* 'door'.

(3) Québec-Labrador Inuktitut (Dorais 1988: 27)
illu-up *ukkua-**nga***
house-GEN door-3SG.POSS
'the door of the house'

It should be stressed that the term *possessive* is used here, as it is generally used by linguists, to denote a range of meanings that is broader than that associated with the word *possession* in everyday English. While it includes a relationship of ownership or possession in the everyday sense, it also includes kinship relations (as in *my father*), part-whole relations (as in *my hand*), and other relations as in *its population, its foreign minister*, *his first novel*, or *his favourite movie*.

The map shows the distribution of **possessive prefixes**, as in (1a) above, in contrast to **possessive suffixes**, as in (1b). Some languages have both possessive prefixes and possessive suffixes. If neither can be considered primary, then the language is shown on the map as having **both possessive prefixes and possessive suffixes with neither primary**. For example, Squamish (Salishan; British Columbia) employs a combination of prefixes and suffixes, as illustrated by the paradigm in (4).

(4) Squamish (Kuipers 1967: 87)

*ʔ**n**-snəxʷiʼλ*	'my canoe'
ʔə-snəxʷiʼλ	'your (singular) canoe'
*snəxʷiʼλ-**s***	'his canoe'
*snəxʷiʼλ-**čət***	'our canoe'
*ʔ**ə-**snəxʷiʼλ-**i̯ap***	'your (plural) canoe'
*snəxʷiʼλ-s-**u̯it***	'their canoe'

However, if there is a reason to view either prefixes or suffixes as primary, the language is shown according to that primary type. For example, in Lillooet, which like Squamish is a Salishan language of British Columbia, there are possessive suffixes on nouns for five of the six combinations of person and number, but for one (first person singular), a prefix is used, as in (5).

(5) Lillooet (van Eijk 1997: 145)

n-tmixʷ	'my land'	*tmíxʷ-**kaɬ***	'our land'
*tmíxʷ-**su***	'your land'	*tmíxʷ-**laꝑ***	'your (plural) land'
*tmixʷ-**s***	'his land'	*tmíxʷ-**i***	'their land'

Because the affixes are primarily suffixes, Lillooet is treated as a language in which possessive suffixes are primary. Similarly, Maricopa (Yuman; Arizona) employs possessive prefixes that represent the person of the possessor and, if the possessor is plural, possessive suffixes that code the plurality of the possessor. The example in (6a) illustrates a noun with a first-person singular possessor, while (6b) illustrates a noun with a first-person plural possessor.

(6) Maricopa (Gordon 1986: 34)

a. *ʔ-ime*
1.POSS-leg
'my leg(s)'

b. *ʔ-ime-**sh***
1.POSS-leg-PL.POSS
'our legs'

Since possessed nouns in Maricopa can contain a possessive prefix without a possessive suffix, but not a possessive suffix without a possessive prefix, the prefixes are considered primary and Maricopa is shown on the map as having possessive prefixes.

The fourth type of languages shown on the map are those that have **no possessive affixes**, but express pronominal possession by means of separate words modifying the noun, as in English *my hat*. Note that languages of this sort are proportionally underrepresented on the map; they are much more common than their frequency on the map might suggest.

Some languages have what at first sight look like possessive affixes in that they can attach phonologically to nouns, but which on more careful examination turn out to be possessive clitics with a syntactically defined position within noun phrases that under certain circumstances attach to nouns, but in other circumstances attach to a nominal modifier. For example, while (7a) from Tukang Besi (Austronesian; Sulawesi, Indonesia) looks like it contains a possessive suffix, the example in (7b) shows that the possessive marker can attach to a postnominal modifier, showing that it is a clitic.

(7) Tukang Besi (Donohue 1999: 73)

a. *te* *kene=su*
CASE friend=1SG.POSS
'my friends'

b. *te* *wunua* *molengo=su*
CASE house old=1SG.POSS
'my old house'

Such clitics are not treated as possessive affixes for the purposes of this map, and Tukang Besi is thus shown as a language lacking possessive affixes.

In many languages, possessive affixes are used only for certain possessive relationships, separate possessive pronominal words being used for other relationships. It is common, for example, for possessive affixes to be used for instances of **inalienable possession**, where the possessive relationship is an inherent or permanent one, but not for alienable possession, where the possessive relationship is in principle a temporary one (see Chapter 58). The two most common subtypes of inalienable possessive relationships are kinship and body-part relations. For example, in Loniu, possessive suffixes are used for inalienable possession, as illustrated in (1b) above, but not for alienable possession, where a pronominal possessor is expressed by a possessive particle followed by an independent pronoun, as in (8).

(8) Loniu (Hamel 1994: 44)
pʷɛlɛyah a yo
parrotfish POSS 1SG
'my parrotfish'

In Ungarinjin (Wororan; northwestern Australia), there are two different sets of possessive affixes, a set of prefixes used with a subset of body-part terms, as in (9a), and a set of suffixes used with kinship terms, as in (9b); alienably possessed nouns do not employ possessive affixes.

(9) Ungarinjin (Rumsey 1982: 43, 47)
a. *ŋiy-embularu*
1SG.POSS-foot
'my foot'
b. *gayi-**ŋi***
grandmother-**1SG.POSS**
'my grandmother'

In Paumarí (Arauan; Brazil), there is one set of possessive affixes for body parts and a different set for kinship terms and alienable possession (Chapman and Derbyshire 1991: 256). In some languages, the use of possessive affixes is restricted to kinship terms; an example is Mparntwe Arrernte (Pama-Nyungan; Northern Territory, Australia; Wilkins 1989: 133). Conversely, in Tauya (Madang, Trans-New Guinea; Papua New Guinea), possessive affixes are only used with body-part terms, separate words being used with other nouns, including kinship terms (MacDonald 1990: 129, 131).

Many languages extend the use of inalienable possession beyond kinship terms and body parts, and often the boundary between alienable and inalienable possession is apparently partly lexicalized. For example, in Kiribati (Oceanic, Austronesian; Kiribati, Pacific), one of the words for 'house' takes the suffix used with inalienable possession, while another word for 'house' takes the separate possessor word used with alienable possession that precedes the possessed noun, as illustrated in (10).

(10) Kiribati (Groves et al. 1985: 49)
a. *mweenga-**ra***
house-**1PL.POSS**
'our house'
b. *ara auti*
1PL.POSS house
'our house'

While there may be some semantic difference between these two nouns that would explain this difference, Groves et al. (1985: 49) give no indication of such and simply describe the example in (10b) as "exceptional".

While the distinction between alienable and inalienable possession is relevant in some languages for whether or not possessive affixes are used, there are other languages that employ possessive affixes for both types of possession, but where the particular nature of the morphological construction varies according to whether the possession is alienable or inalienable. For example, Alune (Austronesian; Ceram, Indonesia) employs possessive suffixes for inalienable possession, but possessive prefixes for alienable possession (Niggemeyer 1951–2: 62).

There are also languages where the use of possessive affixes as opposed to separate possessive pronominal words depends on the particular person-number combination. In Nivkh (isolate; Sakhalin Island, Russia), there are possessive prefixes only for singular possessors; independent pronouns are used for dual and plural possessors (Gruzdeva 1998: 28). Similarly, in Wolof (Atlantic, Niger-Congo; Senegal and Gambia), the third-person singular possessive is realized by a suffix on the noun, while other person-number combinations involve a separate possessive word preceding the noun (Sauvageot 1965: 94). As long as a language regularly employs possessive affixes under some set of circumstances, it is shown on the map according to the position of those affixes.

2 Geographical distribution

The map shows what is perhaps the clearest apparent example in this atlas of an Old World–New World split in the distribution of the two types of possessive affixes: while possessive suffixes are the primary type in the Old World, possessive prefixes are primary in the New World. One must be wary, however, of concluding that this pattern reflects large-scale areal (or genealogical) relationships; to some extent, it may be coincidental.

In Africa, possessive prefixes are relatively uncommon, though instances of them are scattered over the continent. In Europe and Asia, suffixes again predominate, the largest set of exceptions being Tibeto-Burman languages in an area centred in north-east India, all of which employ prefixes if they have possessive affixes. The other genealogical groups containing languages with prefixes in Europe and Asia are North-West Caucasian, Jarawa (an Andamanese language and four isolates: Burushaski, Ket, Nivkh, and Ainu. The Austronesian languages with possessive affixes are overwhelmingly suffixing.

Perhaps the largest overall exception to the appearance of an Old World–New World split is provided by non-Austronesian languages of the region in and around New Guinea. In terms of numbers of genera (linguistic groups comparable in time depth to the subfamilies of Indo-European; see Introduction to Genealogical Language List) containing languages of the two affix types, possessive prefixes strongly outnumber possessive suffixes in this region, by nineteen genera to seven. When one looks at the map of New Guinea, the two affix types look comparable in frequency, with prefixes more common to the west and suffixes more common to the east, but the distribution is largely predictable genealogically: all Austronesian languages of the region with possessive affixes employ suffixes. This predictability from genealogical classification extends eastward: on Bougainville and in the Solomon Islands the Austronesian languages with possessive affixes employ suffixes while the non-Austronesian languages employ prefixes.

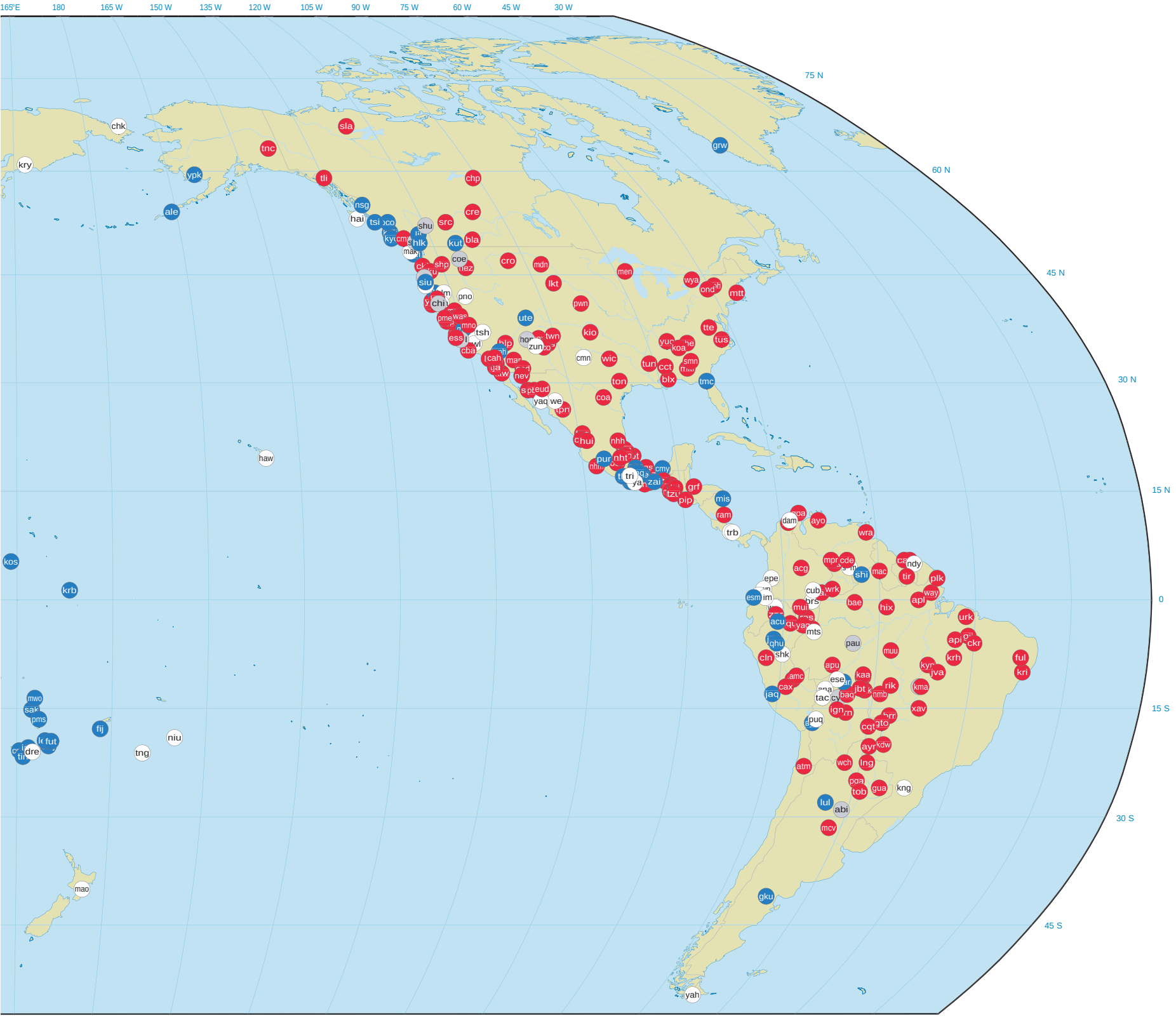
57 Position of Pronominal Possessive Affixes

MATTHEW S. DRYER

Possessive prefixes [238]
Possessive suffixes [330]
Prefixes and suffixes [24]
No possessive affixes [203]

eastern Africa

New Guinea and northern Australia



western North America

Mesoamerica

58 Obligatory Possessive Inflection

JOHANNA NICHOLS AND BALTHASAR BICKEL

1 Bound "inalienables"

In many languages with head-marked possession (see Chapter 24) some nouns obligatorily require possessive inflection and cannot be used alone. For example, the nouns illustrated in (1) and (2) from Navajo (Athabaskan; New Mexico and Arizona) and Acoma (Keresan; New Mexico) cannot stand alone and require possessive inflection.

(1) Navajo (Young and Morgan 1987: 3)
-be' 'milk'
bi-be' 'her milk'

(2) Acoma (Miller 1965: 160)
-jáẓa 'horn'
záça 'his horn'

These are often called *bound nouns*; here we use the more cumbersome but more precise term **obligatorily possessed nouns**. The opposition of obligatorily to optionally possessed nouns is often referred to as a distinction between "alienable" and "inalienable" possession (terminology that is also used in several other senses, one of them mentioned below).

In practical dictionaries obligatorily possessed nouns are usually cited in one or another of the possessive forms (see, for example, Young and Morgan 1987: 2–3 for discussion of which possessive form is used for which nouns in the Navajo dictionary: third person singular for kin terms, unspecified/inanimate possessor for nouns such as body parts that might belong to an animal). In scientific works they are often cited in stem form with a hyphen (e.g. Miller 1965 for Acoma, citing the examples in (3) below). The main map shows languages for which obligatorily possessed nouns are reported.

1.	Obligatorily possessed nouns exist	43
2.	No obligatorily possessed nouns	201
	total	244

Note that obligatorily possessed nouns are not simply nouns that are usually or prototypically used with a possessor, like English *hand* or *sister*; they are words for which an inflectional category of possession is obligatorily present, and which cannot be self-standing. We found examples of obligatorily possessed nouns only in languages where possession is head-marked in the form of affixal morphology.

Languages with obligatorily possessed nouns often provide means of using these nouns independently without a possessor. Quite often the regular possessive inflectional paradigm includes an "indefinite" or "unspecified" possessor category 'someone's, something's', e.g. Navajo *'a-be* [UNSP-milk] 'something's milk, someone's milk, some animal's milk', which is more or less equivalent to English *milk* by itself. Other languages have derivational processes that turn obligatorily possessed nouns (or some of them) into nouns that need not be possessed, which we will call **optionally possessed nouns** (they could also be called *free*, or *self-standing*, in opposition to *bound*). An example is Acoma, illustrated in (3) below, in which 38 of the many obligatorily possessed nouns form corresponding optionally possessed nouns.

(3) Acoma (Miller 1965: 160–1)

bound stem	derived free noun	
-áwíc'i	*wíc'ini*	'chest'
-ánásgai	*nasgâini*	'head'
-'áadák'a	*háadák'ani*	'heel'
-'ásdi'i	*h'aasdí'ini*	'foot'
-jaẓa	*háça̧ni*	'horn'

(All of these pairs, including 'horn', involve regular morphophonemic alternation.) Commonly for derivationally produced optional possessibility, and sometimes for inflectionally produced optional possessibility, the secondary free noun can then itself be inflected for possession, producing a semantic difference that is sometimes described in the literature as an opposition of "alienable" to "inalienable" possession. Examples are in (4) and (5): the first line of each example is the "inalienable" and the third line is the "alienable".

(4) Navajo (Young and Morgan 1987: 3)

possessed:	*bi-be'*	'her milk (from her own breasts)'
unspecified possessor:	*'a-be'*	'something's milk'
possessed unspecified:	*be-'a-be'*	'her (store-bought) milk'

(5) Acoma (Miller 1965: 160–1)

possessed:	*záça*	'his horn' (e.g. a stag's own horn)
derived free noun:	*háça̧ni*	'horn', 'a horn'
possessed free noun:	*k'aháça̧ni*	'his horn' (e.g. a horn belonging to a person)

Obligatorily possessed nouns are found chiefly in the Americas, where they are very common. "Alienability" oppositions like those in (4)–(5) are reported for a number of American languages, chiefly for body-part nouns and (in our sample) only in languages with the inflectional and/or derivational means to produce optionally possessed counterparts to obligatorily possessed nouns.

2 Possessive nouns

The typological antithesis to obligatorily possessed nouns is nouns that cannot take the head-marked possessive inflection available in the language. There are a number of languages with such nonpossessible nouns in our sample, and most of these languages have nouns conventionalized or grammatically specialized for use in apposition to the possessed noun and bearing the inflectional possessive marking which the head noun cannot bear. For instance, in Guaraní (Tupi-Guarani; Paraguay), various nouns including animal names cannot take the usual head-marked possessive inflection. In the event that possession needs to be expressed for such a noun, one or another abstract or generic noun is put in apposition to the semantically possessed nonpossessible, and that abstract or generic noun is inflected for possession:

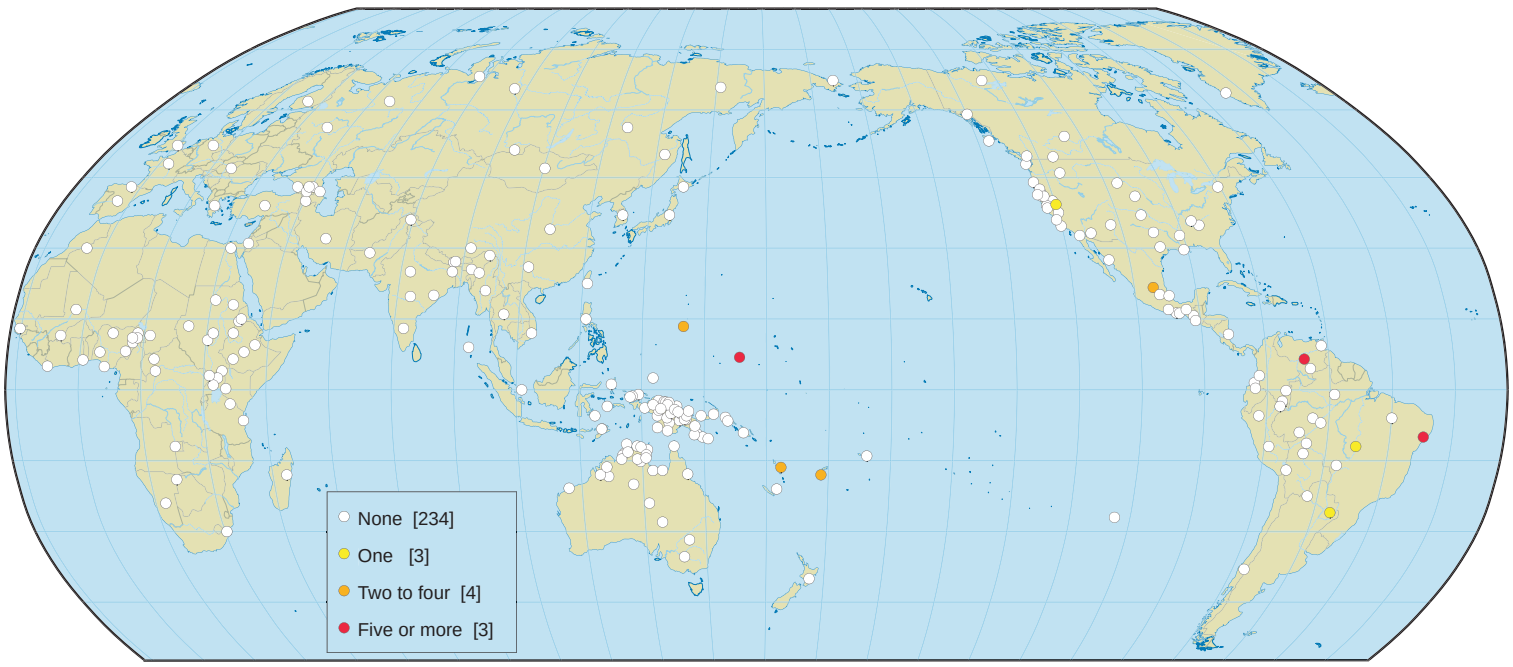
(6) Guaraní (Jensen 1998: 503)

mamaz	*r-eimaw*	*zapukaz*
mother	POSS.LINK-pet	chicken

'mother's chicken'

The head noun 'chicken' is grammatically nonpossessible, though as the example indicates its referent can be owned, showing that nonpossessibility is a grammatical rather than a purely semantic property. The formally possessed generic noun 'pet' is a more or less conventional choice in this construction. Such nouns will be called **possessive nouns** here.

In most languages with possessive nouns, there is a small and closed set of nouns grammaticalized for such usage, and their semantics amounts to classification of possessive relations (so that possessive nouns are also referred to as **possessive classifiers**). Chichimeca-Jonaz, for instance, has a set of four possessive classifier nouns (themselves obligatorily possessed) used with nonpossessibles and classifying them as food, clothing, animals, and things (Lastra de Suárez 1984: 25). A few systems are more elaborate. For Kariri (Macro-Ge; Brazil), Rodrigues (1999: 191) describes a set



Map 58A Number of Possessive Nouns

of twelve generic nouns that classify possession by the origin of the possessive relationship: found, raised from seed, received as a gift, received in a regular partition of goods, etc. Panare (Carib; Venezuela) has twenty-one possessive classifiers which mostly reflect the shape, function, purpose, etc. of the possessed noun (Aikhenvald 2000: 128, citing Carlson and Payne 1989: 19). Maricopa (Yuman; Arizona) has a two-classifier system (in addition to two "inalienable" classes), one for animals and one for all other nouns. Luiseño (Uto-Aztecan; California) has only one classifier, used for animals, as well as obligatorily possessed nouns and nouns which may be possessed or not (Kroeber and Grace 1960: 82–3). The possessive classifier systems of the Americas almost always include a dedicated classifier for animals.

Another language which, like Guaraní, uses more or less conventionalized generic nouns for possession of nonpossessibles is Washo (isolate; Nevada), which also has a derivational prefix that creates derived possessibles from other nonpossessibles (Jacobsen 1964: 408, 468). Nonpossessible nouns are also reported for Hixkaryana (Carib; Brazil), the Arawakan family of South America, and Tzutujil (Mayan; Guatemala) without explicit indication of whether conventionalized generic nouns can be used to make them possessible. The nonpossessible nouns in all of these languages seem to be smallish, closed, or specifiable classes, and they almost always include names of animals. In Kiowa (Kiowa-Tanoan; Oklahoma etc.) body-part nouns cannot take possessive marking but require their possessor to appear as an external possessor.

Appositive possessive nouns are especially well studied for the Oceanic languages (see Lichtenberk 1983b). For example, in Paamese (Vanuatu), there are four possessive nouns which classify the possessive relation as intended for drinking, intended for eating, legally constituted ownership, and use or manipulation. That the choice is not a matter of gender or declension classification by the head noun is shown by examples like (7), in which the same head noun *ani* 'green coconut' can be used with all four possessive nouns depending on the purpose or social basis of the possession.

(7) Paamese (Crowley 1982: 60)
The interlinears 'POT[able]', 'ED[ible]', 'LEG[ally constituted ownership]', and 'MAN[ipulable]' label Crowley's four semantic types.

a. *ani* *ma-k*
coconut POSS.POT-1SG.POSS
'my green coconut' (speaker intends to drink the juice)

b. *ani* *aa-k*
coconut POSS.ED-1SG.POSS
'my green coconut' (speaker intends to eat the meat)

c. *ani* *sa-k*
coconut POSS.LEG-1SG.POSS
'my green coconut' (e.g. growing in speaker's plantation)

d. *ani* *ona-k*
coconut POSS.MAN-1SG.POSS
'my green coconut' (speaker will use it as an implement, e.g. to flatten something)

Fijian has a semantically and formally similar three-way opposition (Dixon 1988: 120–8). Pohnpeian (Micronesia) has a large and possibly open set of possessive classifiers (Rehg 1981: 179–85 lists 21 important ones).

The inset map shows the numbers of possessive nouns in the ten languages in which we found them.

1.	None reported	234
2.	One	3
3.	Two to four	4
4.	Five or more	3
	total	244

Values of Map 58A Numbers of Possessive Nouns

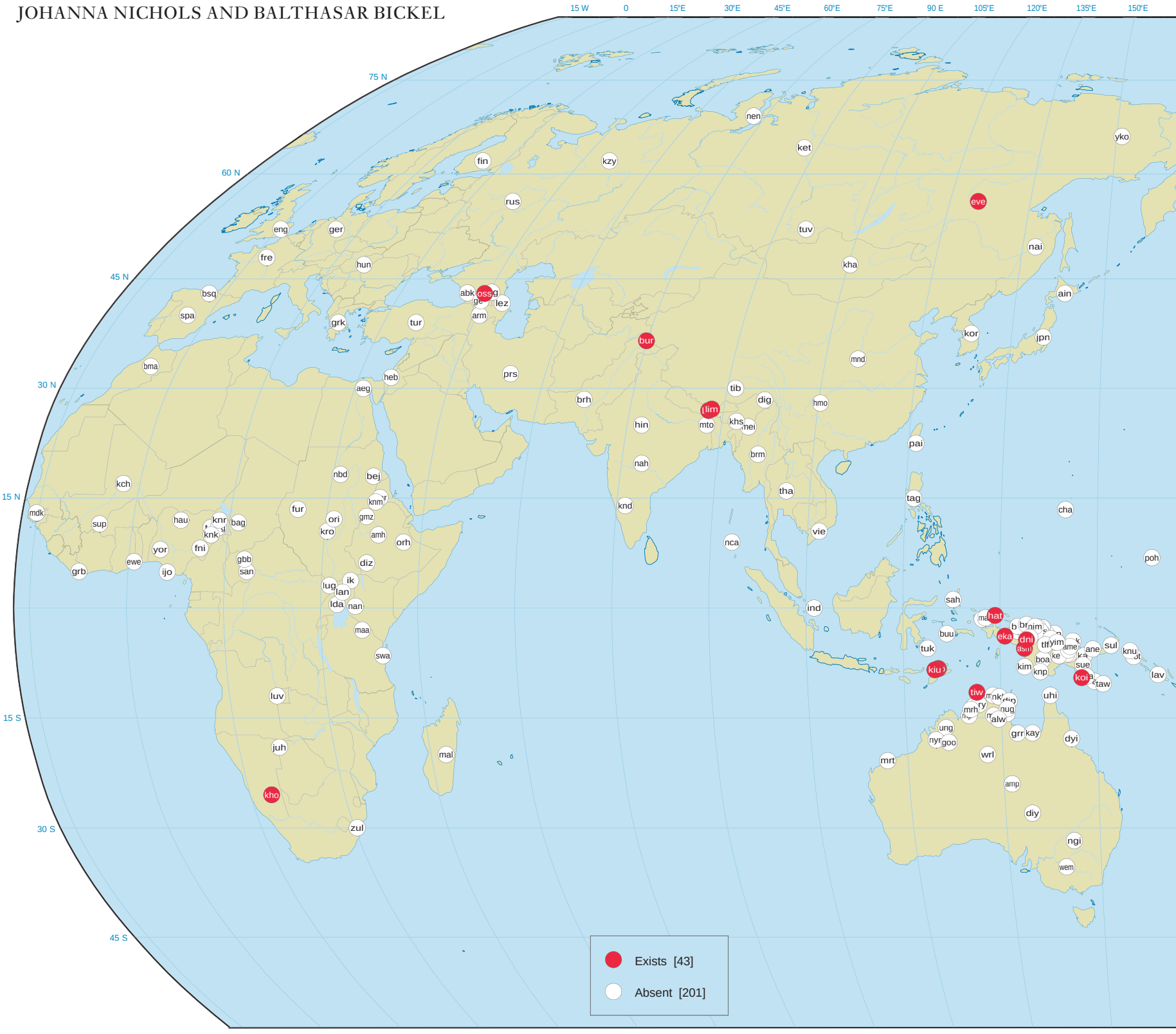
(The numbers of possessives actually attested are one, three, four, twelve, and twenty-one.) Included are the possessive classifier systems like those of Paamese, Fijian, Chichimeca-Jonaz, etc., as well as the use of generic nouns with nonpossessibles. Classificatory possessive noun systems are a Pacific Rim feature (exclusively so in our sample; Aikhenvald (2000: 130) mentions one African language), and are particularly well developed at the far extremes of the Pacific Rim: in the outer Pacific islands (e.g. Pohnpeian, mentioned above) and South America, especially Amazonia.

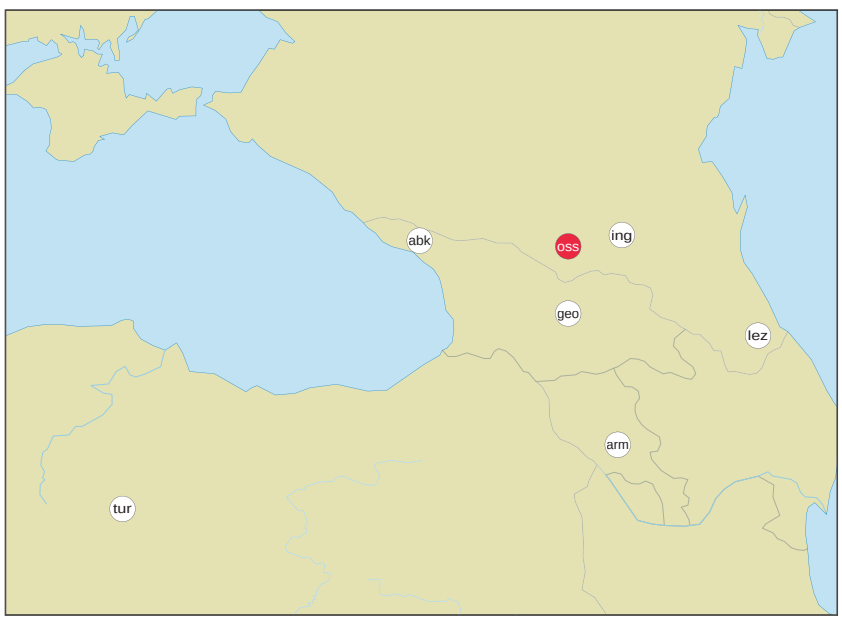
Acknowledgements

This research was supported in part by US National Science Foundation Grant No. 96-16448 (Nichols, P.I.), Swiss National Science Foundation Grants No. 08210-053455 and 610-062717 (Bickel, P.I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, David A. Peterson, and Suzanne Wilhite for help with data collection.

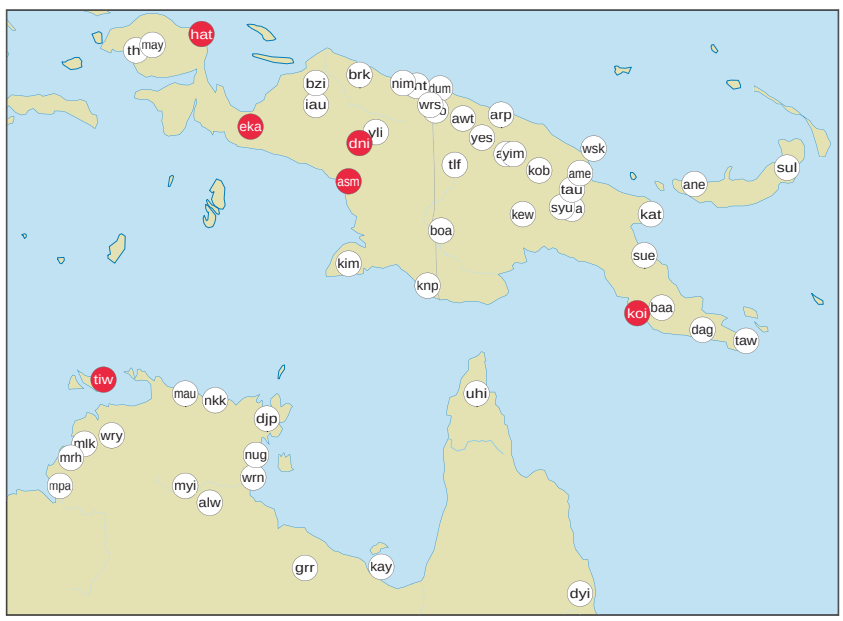
58 Obligatory Possessive Inflection

JOHANNA NICHOLS AND BALTHASAR BICKEL

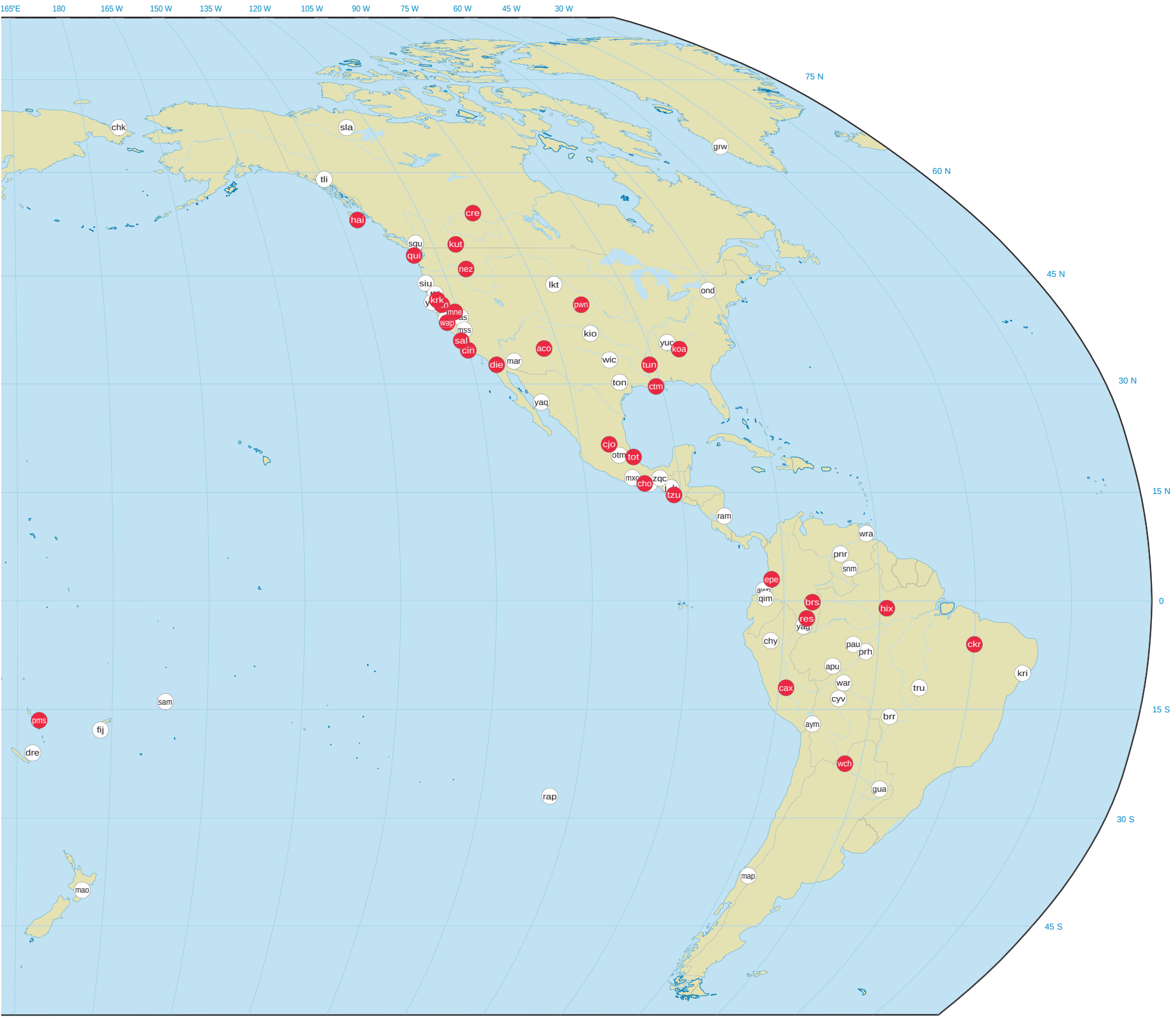


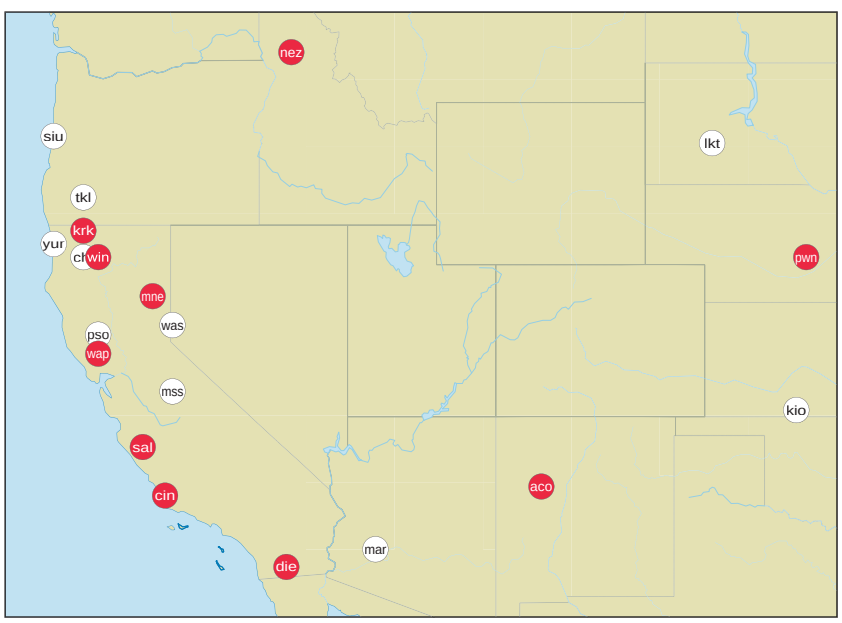


Caucasus and adjacent areas

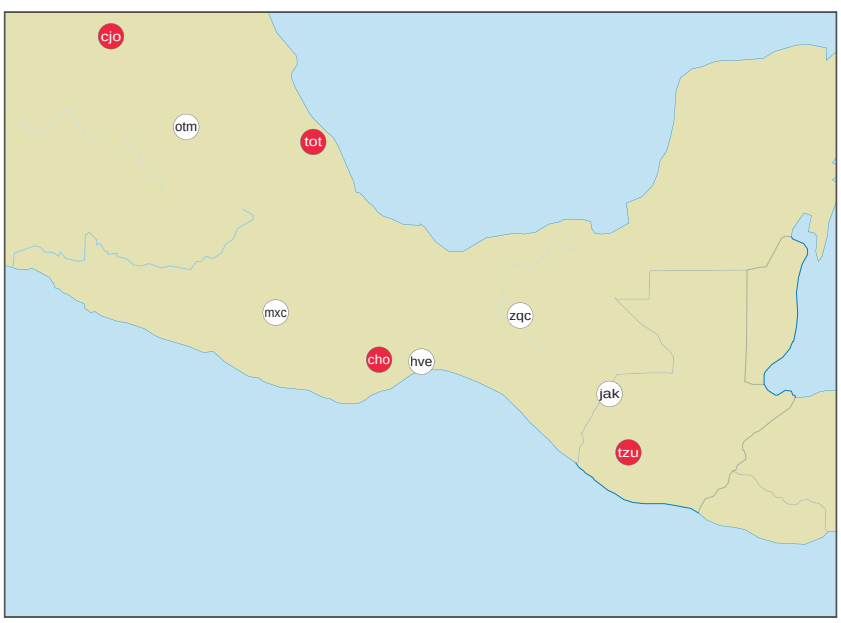


New Guinea and northern Australia





western United States



southern Mexico

59 Possessive Classification

JOHANNA NICHOLS AND BALTHASAR BICKEL

1 Possessive classes

Many languages have more than one way of forming possessive noun phrases. English, for example, has a choice between the "Saxon genitive" (*Mary's house, the director's office, the day's end*) and the "Norman genitive" (*the office of the director, the end of the month*), conditioned in part by the part of speech and semantics of the possessor but with considerable overlap and stylistic variation.

In contrast, many languages have an opposition of two (or sometimes more) forms of possessive marking whose choice is conditioned not by semantics or style, as in English, but **lexically**; and conditioned not by properties of the possessor but **by the possessed noun**, i.e. by the head noun in the construction. (The possessed nouns in the English examples above are *house*, *office*, and *end*.) Consider the examples in (1) from Mesa Grande Diegueño (Yuman; California), where the noun 'mother' takes the simple prefix *ʔ-* 'my' while 'house' takes the compound prefix *ʔə-nʸ-*.

(1) Mesa Grande Diegueño (Langdon 1970: 143, 145)

a. *ʔ-ətalʸ*
1SG-mother
'my mother'

b. *ʔə-nʸ-əwaː*
1SG-ALIENABLE-house
'my house'

This contrast of two formal types of possession, determined by the possessed noun, is **possessive classification**. We use the terms *possessive* and *possession*, as is traditional, to refer to all kinds of adnominal constructions regardless of whether the semantics is literal possession. We use the term *classification* because the marking of possession in examples like (1) divides, or classifies, the nouns of Mesa Grande Diegueño into two sets: those that behave grammatically in possessive NPs like 'mother' and those that behave like 'house'.

Though in many languages the two kinds of possession use the same kind of morphology (as in Mesa Grande Diegueño, where both kinds use prefixes on the possessed noun), a number of languages use entirely different kinds of morphology in the different possessive constructions. An example comes from Warndarang (Maran; Northern Territory, Australia). Here kin terms and certain others take possessive prefixes as in (2a), while other nouns take a separate possessive pronoun in the genitive case as in (2b).

(2) Warndarang (Australia: Heath 1980: 28–9, 34)

a. *nga-baba*
1-father
'my father', 'our father'

b. *wuradburru* *ngini*
camp 1SG.GEN
'my camp'

In a number of languages, one form of possession involves simple juxtaposition and no marking while the other involves overt marking, as in (3) from Tiwi (isolate; Northern Territory, Australia).

(3) Tiwi (Osborne 1974: 74–5)

a. *jərəkəpai* *tuwaɹa*
crocodile tail
'(the/a) crocodile's tail'

b. *jərəkəpai* *ŋara* *tuwaɹa*
crocodile he tail
'(the/a) crocodile's tail'

In Tiwi, either construction is possible for most nouns, but kin terms take only juxtaposition (as in 3a).

The distinctions shown in all these examples are classification because they are lexically determined by the possessed noun and result in an overt difference in the marking of possession. We consider only classes that are lexically conditioned, and have excluded any that are phonologically conditioned (or appear to be amenable to a phonological analysis). We note that, while lexical classifications are most often equipollent (one class of nouns takes possessive marker A and another takes marker B), they can also be privative (as in the Tiwi example just above: one class generally takes A but can sometimes or occasionally take B, perhaps with a semantic opposition; the other class takes only B but never A). Even if there is a semantic opposition available to one class, the overall classification is still lexical.

2 How many classes?

Most languages with possessive classification have a binary opposition like those illustrated above. Sometimes the classification is more elaborate. Nez Perce (Sahaptian; northwestern United States), for example, has three possessive classes: two different possessive prefix paradigms for different kin terms, and a preposed possessor in the genitive case for other nouns:

(4) Nez Perce (Rude 1985: 91, 97; Aoki 1994: 986; see also Aoki 1970: 50)

a. *na'-tóot*
1SG-father
'my father' (prefix paradigm 1)

b. *'in'm-éːks*
1SG-man's.sister
'my sister' (man speaking) (prefix paradigm 2)

c. *'íi-nim* *titóoqan*
1SG-GEN people
'my people'

Burushaski (isolate; Kashmir, India or Pakistan) has four classes of possessive prefixation defined by allomorphy of the possessive prefixes. In (5), classes (a), (c), and (d) have different prefix shapes and (b) causes a stress shift.

(5) Burushaski (Berger 1998: 44, 91)
[(d) 1SG example extracted from text, 2SG constructed]

	1SG	2SG	
a.	*a-yáṭis*	*gu-yáṭis*	'head'
b.	*á-l´cin*	*gú-l´cin*	'eye'
c.	*á-miṣ*	*gó-miṣ*	'finger'
d.	*áa-ṣki*	*góo-ṣki*	'head of bed'

A few languages have complex systems of possessive classification, with ten or more classes. Amele (Madang; Papua New Guinea), for instance, has thirty-one different classes based on the allomorphy of the possessive suffix, as well as an open class which uses a postposition, for a total of thirty-two classes.

(6) Amele (Roberts 1987: 139)

a. *ija* *na* *jo*
1SG of house
'my house' (postposition)

b. *ija* *co-ni*
1SG mouth-1SG
'my mouth' (suffixal)

The following examples illustrate a few of the Amele possessive suffixal classes. The examples chosen all have stem-final *-e* to show that the endings are not phonologically conditioned in any obvious way.

(7) Selected Amele nouns taking different possessive suffixes (Roberts 1987: 172–4)

1st person	2nd person	3rd person	
aide-ni	*aide-n*	*aide-g*	'wife'
ebe-ni	*ebe-in*	*ebe-n*	'hand'
amese-ni	*amese-n*	*amese-c*	'eye socket'
be-ni	*be-n*	*be*	'neck'
qehe-mi	*qehe-m*	*qe(h)i-h*	'side'
mede-mi	*mede-m*	*mede*	'nose'
me-i	*mem-en*	*mem-eg*	'father'

Anêm (isolate; Papua New Guinea) has at least twenty possessive classes based on the allomorphy of the stem-final suffix and of the possessive suffix. Four of these classes are shown in (8).

(8) Anêm (Thurston 1982: 37)

	'water'	'child'	'leg'	'mat'
1SG	*kom-i*	*gi-ŋ-e*	*ti-g-a*	*mîk-d-at*
2SG	*kom-î*	*gi-ŋ-ê*	*ti-g-îr*	*mîk-d-ir*
3SG.M	*kom-u*	*gi-ŋ-o*	*ti-g-î*	*mîk-d-it*
3SG.F	*kom-îm*	*gi-ŋ-êm*	*ti-g-î*	*mîk-d-it*

Chichimeca-Jonaz (Oto-Manguean; Mexico) has twelve (or possibly even more) possessive classes of nouns, where the classes are defined broadly as involving tone changes, prefix-like elements, various internal changes, and combinations of these as well as suppletion. (Each of these classes, or at least some of them, contains both bound and nonbound nouns as these are defined in Chapter 58. In addition, the language has four possessive classifier nouns.) The examples in (9) display this variety.

(9) Chichimeca-Jonaz (Lastra de Suárez 1984: 24–5) (*M* = lenis nasal)

1SG	2SG	
námenʔ	*naménʔ*	'face'
suní	*síni*	'lip'
kútún	*utún*	'neck'
túmbiʔir	*nímbíʔir*	'tail'
nahí	*únho*	'friend'
tásócʔ	*kisócʔ*	'belt'
namá	*éMą*	'carrying rack'
kúndí	*kirí*	'water'
nambá	*úngwa*	'hat'
kúmboʔ	*kibóʔ*	'land'
kaʔá	*kanʔa*	'hand'
masų́	*uniʔí*	'wife'
táta	*úngwæ*	'father'

In Cayuvava (isolate; Bolivia), there are a number of classes—probably at least ten—defined by allomorphy of the possessive affixes (prefix, suffix, or circumfix) (Key 1967: 50).

In these four languages with complex systems (Amele, Anêm, Chichimeca-Jonaz, Cayuvava), possessive classification has become a system of declension classes rather like the declension classes of case inflection in classical Indo-European languages. In Amele the differences between allomorphs are rather straightforward, in Anêm less so, and in Chichimeca-Jonaz and Cayuvava they are great.

Map 59 shows the number of possessive classes in the sample languages.

1. No possessive classification	125
2. Two classes	94
3. Three to five classes	20
4. More than five classes	4
total	243

Binary possessive classification is fairly common everywhere except Eurasia, where it is entirely absent except for tokens in the Pacific Rim and the mountain enclaves (Caucasus, Himalayas). Multiple classes are not particularly common anywhere, though clearly best attested around the Pacific Rim, and the complex systems of ten or more classes are found only around the Pacific Rim. The larger systems of possessive classification in Melanesia are found on or near the coast and chiefly in the north of both Australia and New Guinea. Various kinds of morphology can be involved in possessive classification, as illustrated in the examples above, but all four of the complex systems use head-marking morphology. The complexity and opacity of the Chichimeca-Jonaz and Cayuvava systems would appear to testify to a very great age for the head-marking type in these languages.

3 The semantics of possessive classification

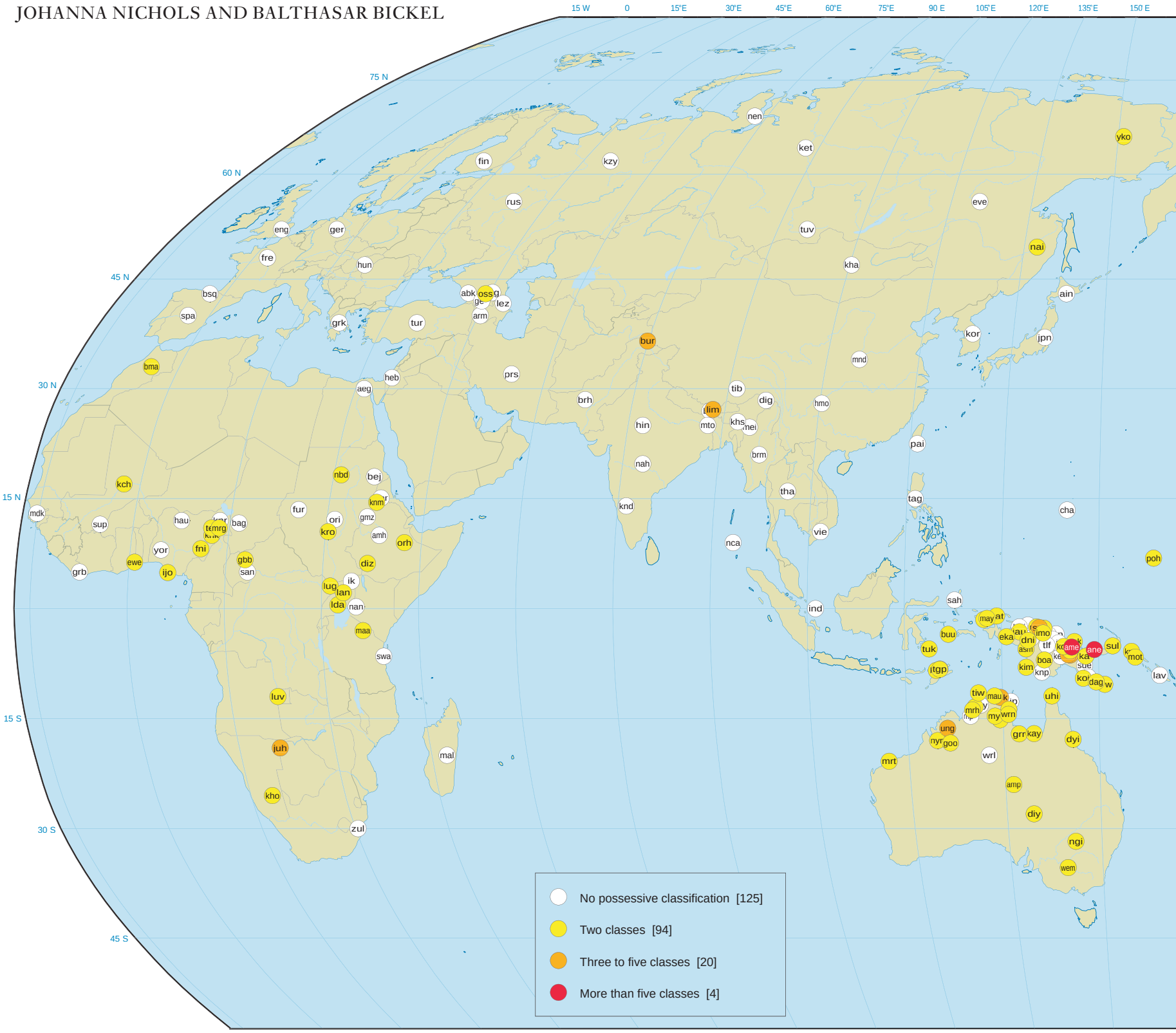
Binary possessive classification is usually called **"alienable/inalienable"** possession in the literature, partly because it often involves examples like those in (1), in which the house but not the mother can be bought, sold, etc., and partly because it is fairly common, especially in the Americas, for one of a binary set of possessive classes to be bound, i.e. obligatorily possessed (see Chapter 58). In fact, though, possessive classification is not a semantic or grammatical category but a purely lexical classification of nouns. Semantics is involved in that there is usually a default or open class and a specified or determinate class with a semantic common denominator shared by most but not all members of the class and also sometimes found in nonmembers. Kin terms often figure in the specified class, and usually several to many of the language's kin terms and a few other nouns take the nondefault possessive form while all others, including a few kin terms, take the default possessive form. In addition, body parts, meronyms (parts in part-whole relationships), topological nouns, and property nouns often figure in nondefault possessive classes.

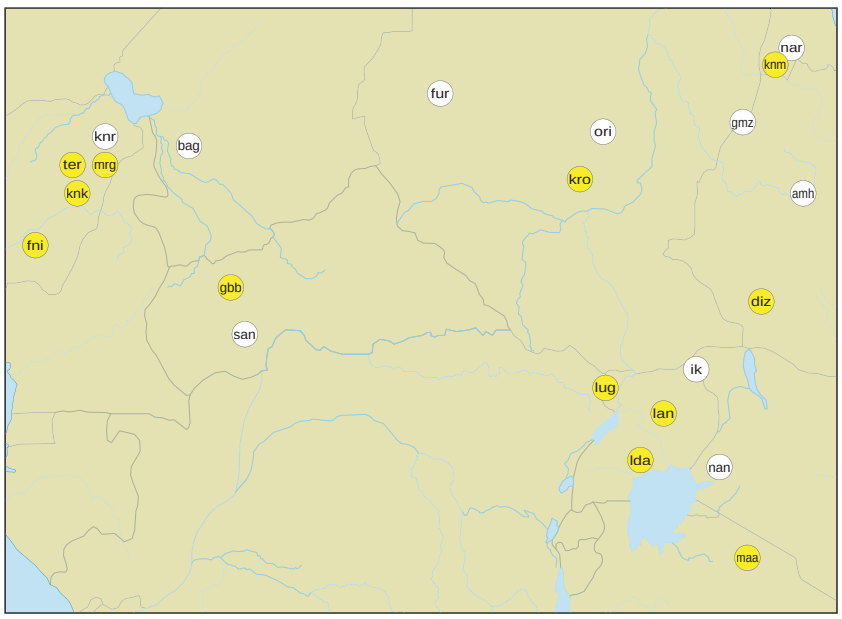
Acknowledgements

Supported by US National Science Foundation Grant 96-16448 (Nichols, P. I.), Swiss National Science Foundation Grants 08210-053455 and 610-062717 (Bickel, P. I.), and the Institute for Slavic, Eurasian, and East European Studies, UC Berkeley. We thank Aimée Lahaussois Bartosik, David A. Peterson, and Suzanne Wilhite for help with data collection.

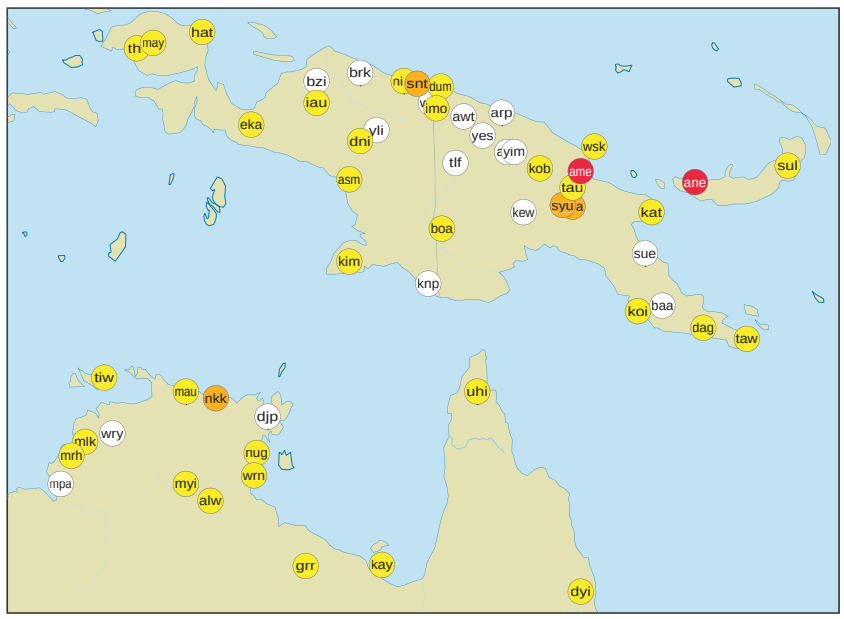
59 Possessive Classification

JOHANNA NICHOLS AND BALTHASAR BICKEL

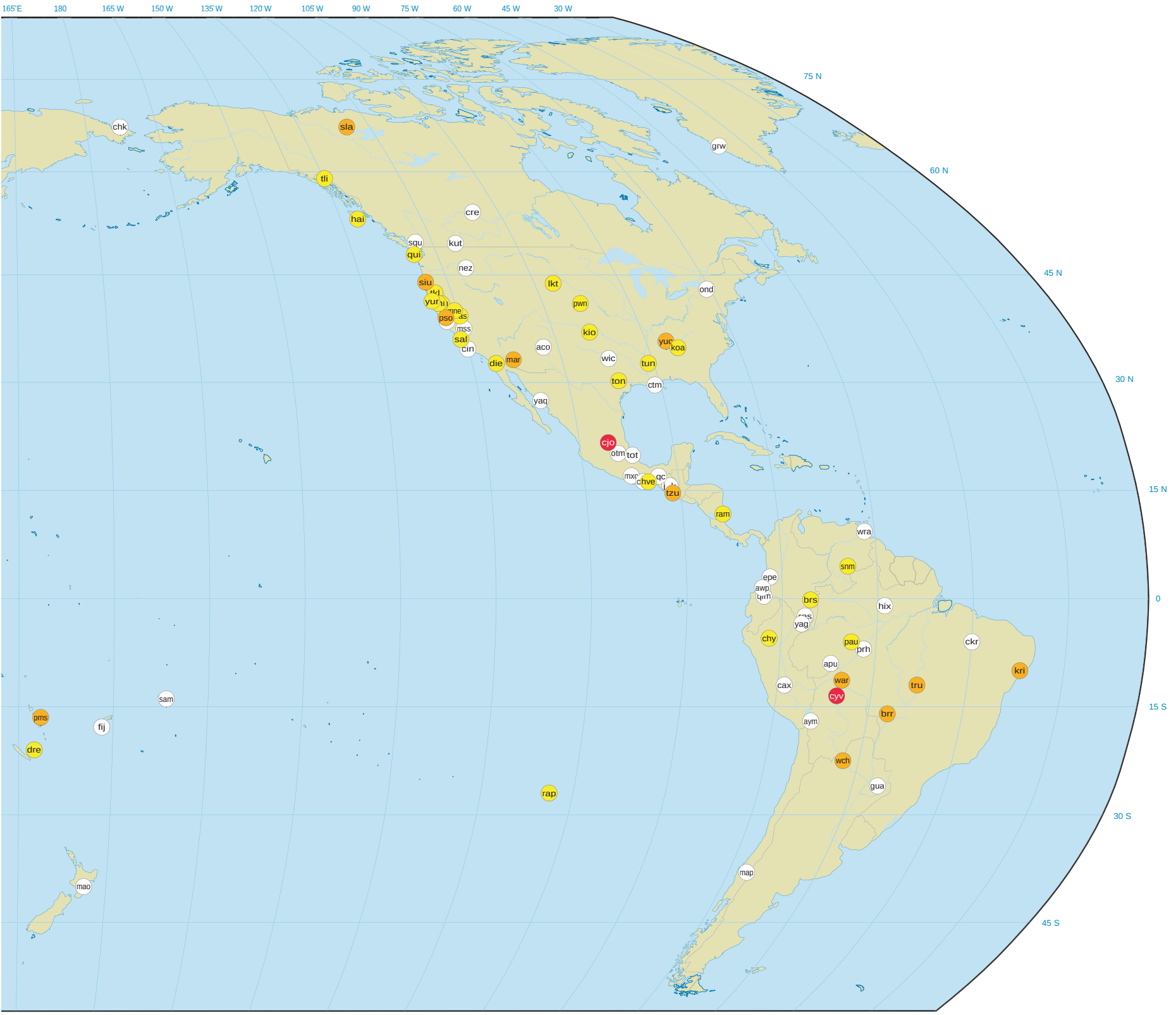


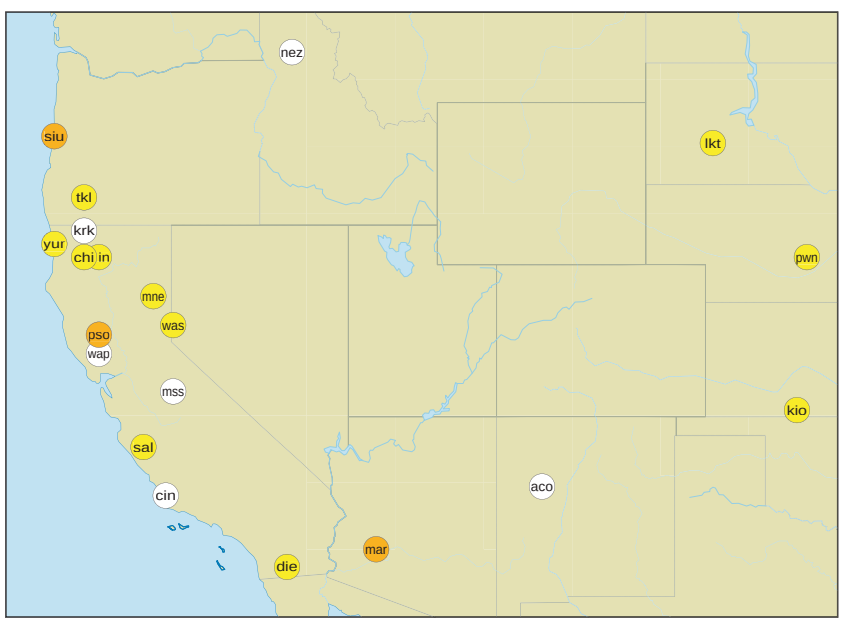


central Africa

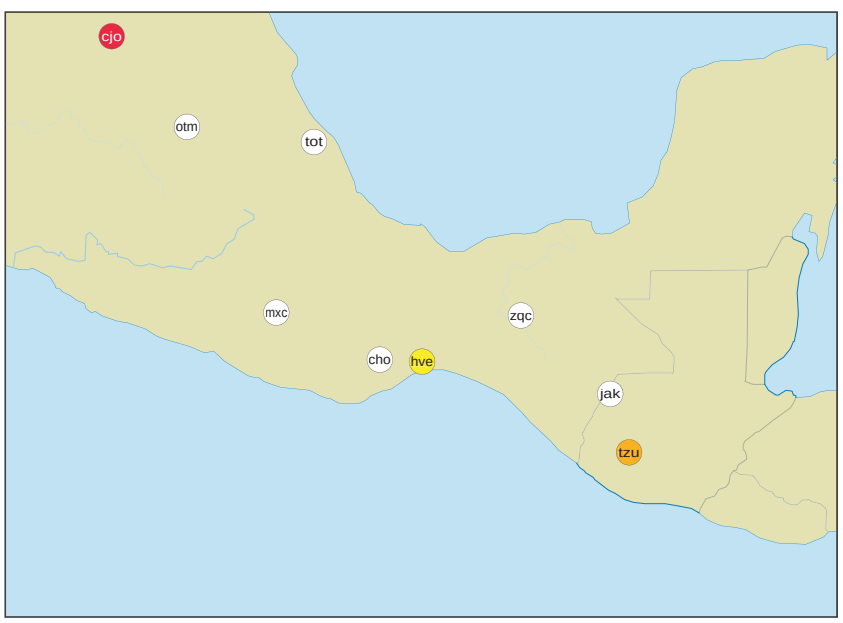


New Guinea and northern Australia





western United States



southern Mexico

60 Genitives, Adjectives, and Relative Clauses

DAVID GIL

1 Introduction

In English, nouns may be modified in a variety of ways, some of which are illustrated in the following example:

(1) a. *John's apple*
b. *red apple*
c. *apple that John bought*

In (1) above, all three NPs consist of a head noun *apple* plus an attributive expression; however, the attributive expression is of a different semantic type in each example: in (1a) it denotes an alienable possessor *John's*, in (1b) a colour property *red*, and in (1c) an activity *that John bought* with respect to which the head noun has the role of patient.

In English, as evident in (1), these three semantic functions are coded by means of three different morphosyntactic constructions: in (1a) the attributive expression precedes the head and is marked with the possessive enclitic *'s*, in (1b) the attributive expression precedes the head but is unmarked, and in (1c) the attributive expression follows the head and is marked with the complementizer *that*. Based on paradigms such as the above in English and in other languages, grammarians distinguish between three different types of constructions, **genitives**, **adjectives** and **relative clauses**, of which the above are typical examples.

However, not all languages possess dedicated morphosyntactic constructions for each of the above semantic functions. In some languages, two of the above functions, or even all three of them, may be expressed by means of a single more general morphosyntactic construction. In such cases, then, one may say that genitives, adjectives, and relative clauses are not fully differentiated to the extent that they are in languages like English.

Thus, languages vary with respect to the degree to which grammatical encoding distinguishes between different semantic types of attribution. The purpose of this map is to depict this cross-linguistic variation, by showing the extent to which different languages distinguish between genitive, adjective, and relative clause constructions.

2 Feature values

The point of departure for this map is provided by the three semantic types of attribution illustrated in (1) above: (a) alienable possession, generally of a human over an inanimate object such as a food item; (b) colour property, or other property such as size or quality; and (c) activity, typically past and punctual, with respect to which the head noun has the role of patient or undergoer. For each language under consideration, the map compares the various ways in which these three semantic functions are grammatically encoded.

Note that a single semantic function may sometimes be encoded by means of more than one grammatical strategy. For example, in English, (1a) may be paraphrased with a postnominal genitive construction *apple of John's*, while (1c) may be alternatively expressed with a relative pronoun as in *apple which John bought*, or without any overt marker as in *apple John bought*. For the purposes of this map, all alternative coding strategies are taken into account, provided that they are deemed to be natural and not "long-winded". The latter proviso is necessary in order to exclude various alternative constructions such as, for example, *apple that John owns* for (1a), and *apple that is red* for (1b).

In order to measure the extent to which a language distinguishes between genitives, adjectives, and relative clauses, the language's maximally general encoding strategy is taken into consideration, and all other more specific strategies are ignored. This makes it possible to rank languages on a three-valued scale, characterizing their attributive constructions as weakly, moderately, or highly differentiated.

1. Weakly differentiated	15
2. Moderately differentiated, with genitives and adjectives collapsed	8
3. Moderately differentiated, with genitives and relative clauses collapsed	2
4. Moderately differentiated, with adjectives and relative clauses collapsed	33
5. Moderately differentiated; other	3
6. Highly differentiated	77
total	138

Languages with **weak differentiation** are those in which there exists a common strategy for the encoding of all three semantic functions. Such languages may also have additional, more specific strategies for the encoding of these functions; however, such strategies are ignored here. One example of a language with weak differentiation is Minangkabau (Sundic, Austronesian; Sumatra, Indonesia), in which all three semantic functions may be expressed by bare postnominal modifiers, without any additional grammatical markers:

(2) Minangkabau (own knowledge)
a. *batiak Kairil*
papaya Kairil
'Kairil's papaya'
b. *batiak kuniang*
papaya yellow
'yellow papaya'
c. *batiak Kairil bali*
papaya Kairil buy
'papaya that Kairil bought'

Another example of a language with weak differentiation is provided by Cantonese, in which all three semantic functions may be expressed by prenominal modifiers marked by the associative particle *ge*^{33}.

(3) Cantonese (Ho Chee Lick, Lee May San, Cassandra Yue Chee Tieng, p. c.)
a. *a*^{33}*-faay*^{55} *ge*^{33} *piŋ*^{11}*-guo*^{35}
Ah Fai ASSOC apple
'Ah Fai's apple'
b. *hooŋ*^{11} *ge*^{33} *piŋ*^{11}*-guo*^{35}
red ASSOC apple
'red apple'
c. *a*^{33}*-faay*^{55} *maai*^{13} *ge*^{33} *piŋ*^{11}*-guo*^{35}
Ah Fai buy ASSOC apple
'apple that Ah Fai bought'

Both Minangkabau and Cantonese have additional, more specific strategies which are not shown here.

Languages with **moderate differentiation** are those that do not have a common strategy for the encoding of all three semantic functions. However, such languages do have one or more strategies which encode two of the three functions to the exclusion of the third. Again, in addition such languages may also have dedicated coding strategies for each of the three semantic functions, which are ignored here. For the purposes of the map, such languages are divided into four subtypes.

In the first subtype, **genitives and adjectives** are collapsed, to the exclusion of relative clauses. An example of such a language is provided by the Västerbotten dialect of Swedish, in which alienable possessors and colour properties may both be expressed by means of a compound construction in which the modifier precedes the head:

(4) Västerbotten Swedish (Christina Alm-Arvius p. c.)

a. *Pelle-äpple*
Pelle-apple
'Pelle's apple'

b. *rö-äpple*
red-apple
'red apple'

In the second subtype, **genitives and relative clauses** are collapsed, to the exclusion of adjectives. An example of such a language is provided by the Jewish Arbel dialect of Aramaic, in which alienable possessors and activities may both be expressed by means of postnominal modifiers marked with *ʔot*. (Note that while adjectives may also be marked similarly, the adjective also agrees with the noun in number and gender; hence this is considered to be a different construction.)

(5) Jewish Arbel Aramaic (Geoffrey Khan p. c.)

a. *qarʕa* *ʔot* *yosef*
pumpkin ASSOC Joseph
'Joseph's pumpkin'

b. *qarʕa* *ʔot* *zwinne* *yosef*
pumpkin ASSOC buy.PST.3SG M Joseph
'pumpkin that Joseph bought'

In the third subtype, **adjectives and relative clauses** are collapsed, to the exclusion of genitives. An example of such a language is provided by Tagalog, in which colour properties and attributive activities may both be expressed by means of modifiers linked to their heads by means of a ligature, which in the following examples assumes the form of an enclitic *=ng*. (Note that while in (6) the modifiers precede their heads, the order is flexible and they could also follow.)

(6) Tagalog (own knowledge)

a. *pula=ng* *mangga*
red=MOD mango
'red mango'

b. *b<in>ili* *ni* *Jojo=ng* *mangga*
<PAT.TOP.PFV>buy PERS.GEN Jojo=MOD mango
'mango that Jojo bought'

In a number of languages, there are two or more strategies which employ identical encoding for different pairs of the three semantic functions. The various possible ways in which this may come about are grouped together in the fourth, **other** subtype of languages with moderate differentiation. One example of such a language is provided by Thai. As shown in (7), alienable possessors and colour properties may both be expressed by means of a common strategy involving bare postnominal modification; in the collapsing of genitives and adjectives Thai thus resembles Västerbotten Swedish. In addition, as indicated in (8), colour properties and activities may both be expressed by means of a common strategy involving postnominal modification with a classifier plus a relative marker; in the collapsing of colour properties and activities Thai thus resembles Tagalog.

(7) Thai (Titima Suthiwan, Uri Tadmor p. c.)

a. *sôm* *kàai*
orange Kai
'Kai's orange'

b. *sôm* *sii* *dææŋ*
orange colour red
'red orange'

(8) Thai

a. *sôm* *lûuk* *tʰîi* *sii* *dææŋ*
orange CLF REL colour red
'red orange'

b. *sôm* *lûuk* *tʰîi* *kàai* *súuu* *maa*
orange CLF REL Kai buy come
'orange that Kai bought'

In other languages of this subtype, different combinations of the three semantic functions may be expressed by means of identical strategies.

Finally, languages with **high differentiation** are those in which there are no common strategies for encoding more than one of the three functions; instead, each of the three functions is encoded by distinct grammatical strategies, genitive, adjective, and relative clause respectively. One example of such a language, discussed above, is English; another, illustrated below, is Abkhaz. In Abkhaz, possessors occur prenominally with the head noun marked with an agreeing possessive form, colour properties occur postnominally in a compound construction, and activities form prenominal relative clauses in which the verb assumes a special relative form.

(9) Abkhaz (Irina Borisovna Ankvad p. c.)

a. *adgur* *jə-ç^{o}a*
Adgur 3SG.POSS-apple
'Adgur's apple'

b. *a-ç^{o}a-q'apš*
DEF-apple-red
'red apple'

c. *adgur* *jaajə-x^{o}a-z* *a-ç^{o}a*
Adgur 3SG.OBJ-buy-PST.REL DEF-apple
'apple that Adgur bought'

3 Geographical distribution

Languages with high differentiation are more common than languages with moderate differentiation, which in turn are more common than languages with low differentiation. And indeed, in most parts of the world, languages with high and moderate differentiation can be found interspersed among each other.

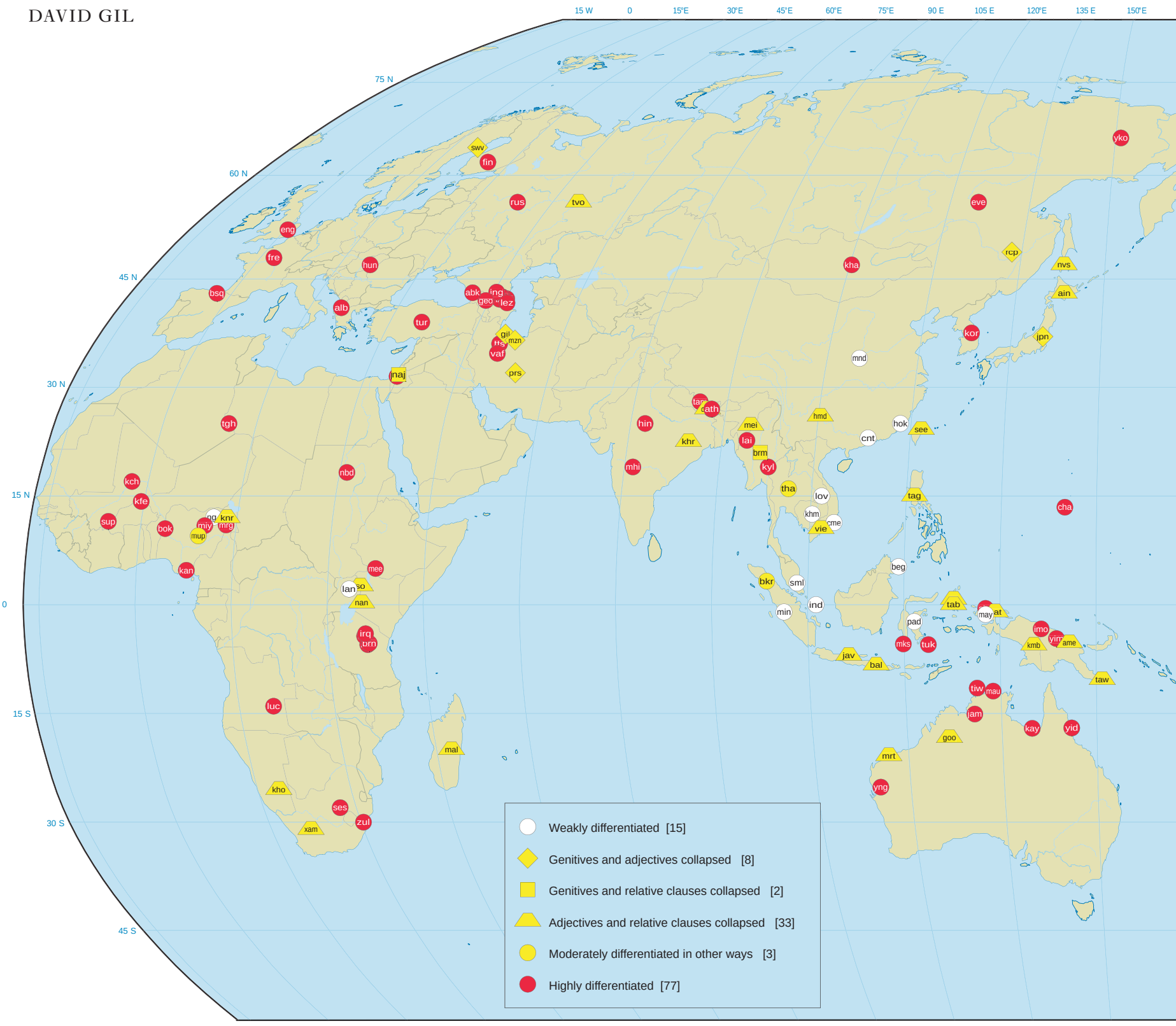
Against this rather homogenous background, two regions stand out as exceptions. First, in a region comprising Europe plus adjacent parts of Asia and Africa, almost all the languages are of high differentiation. Secondly and more strikingly, in a region containing large parts of South-East Asia, almost all the languages are of low differentiation. In fact, this latter region contains almost all the languages of low differentiation in the present sample; outside South-East Asia, such languages are very infrequent.

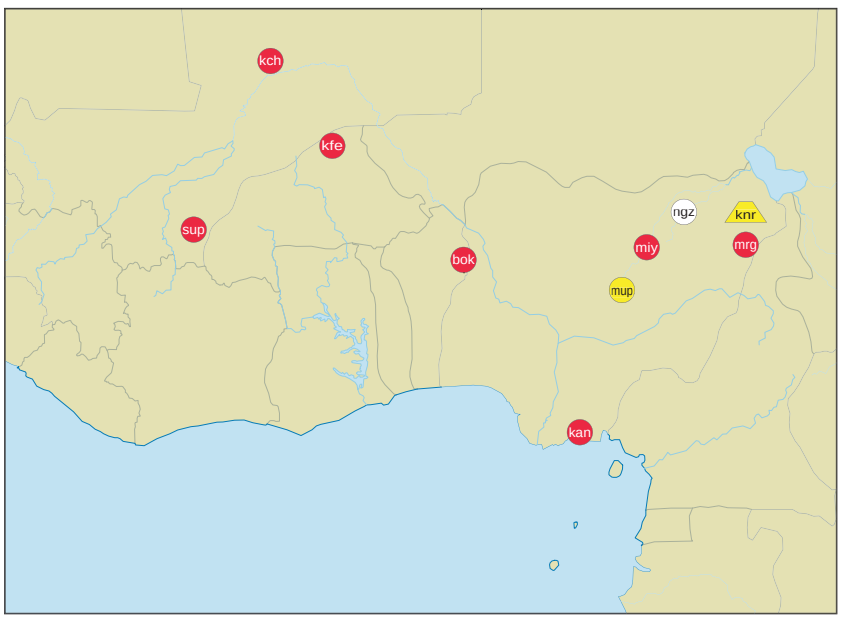
4 Theoretical issues

Many linguists speak European languages in which genitives, adjectives, and relative clauses are highly differentiated. Perhaps because of this, most descriptive and theoretical traditions within current linguistics provide ready-made ways of distinguishing between such constructions. However, such traditions often fail to provide the tools for **not** distinguishing between them, in the case of languages in which genitives, adjectives, and relative clauses are not highly differentiated. Often, linguists coming from European languages will posit distinct genitive, adjective, and relative clause constructions even in languages which seem not to have them. Indeed, when faced with a new language, it is often easier to recognize the existence of some exotic and unfamiliar category than it is to realize that a commonplace and familiar category is in fact absent. However, as argued in Gil (2001), linguists must strive to overcome such predispositions.

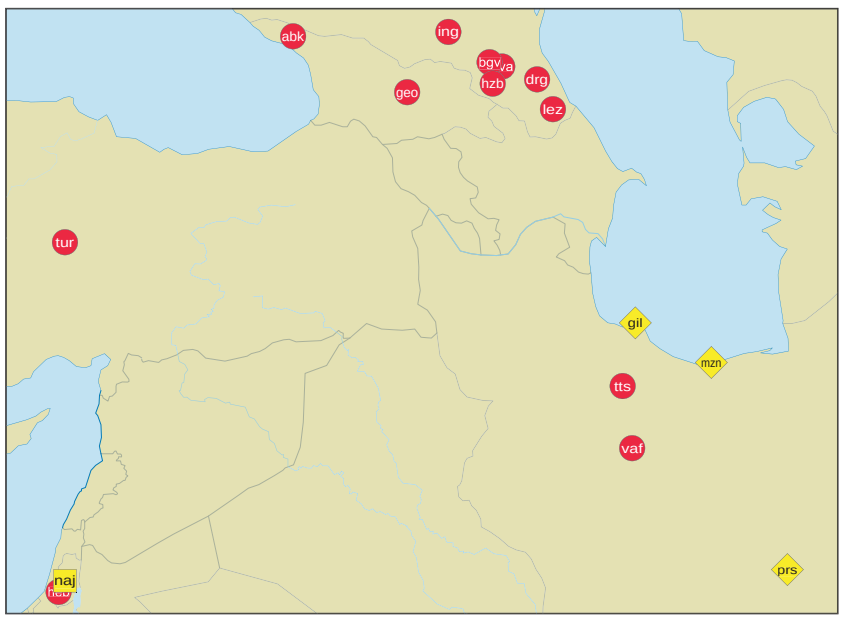
60 Genitives, Adjectives, and Relative Clauses

DAVID GIL

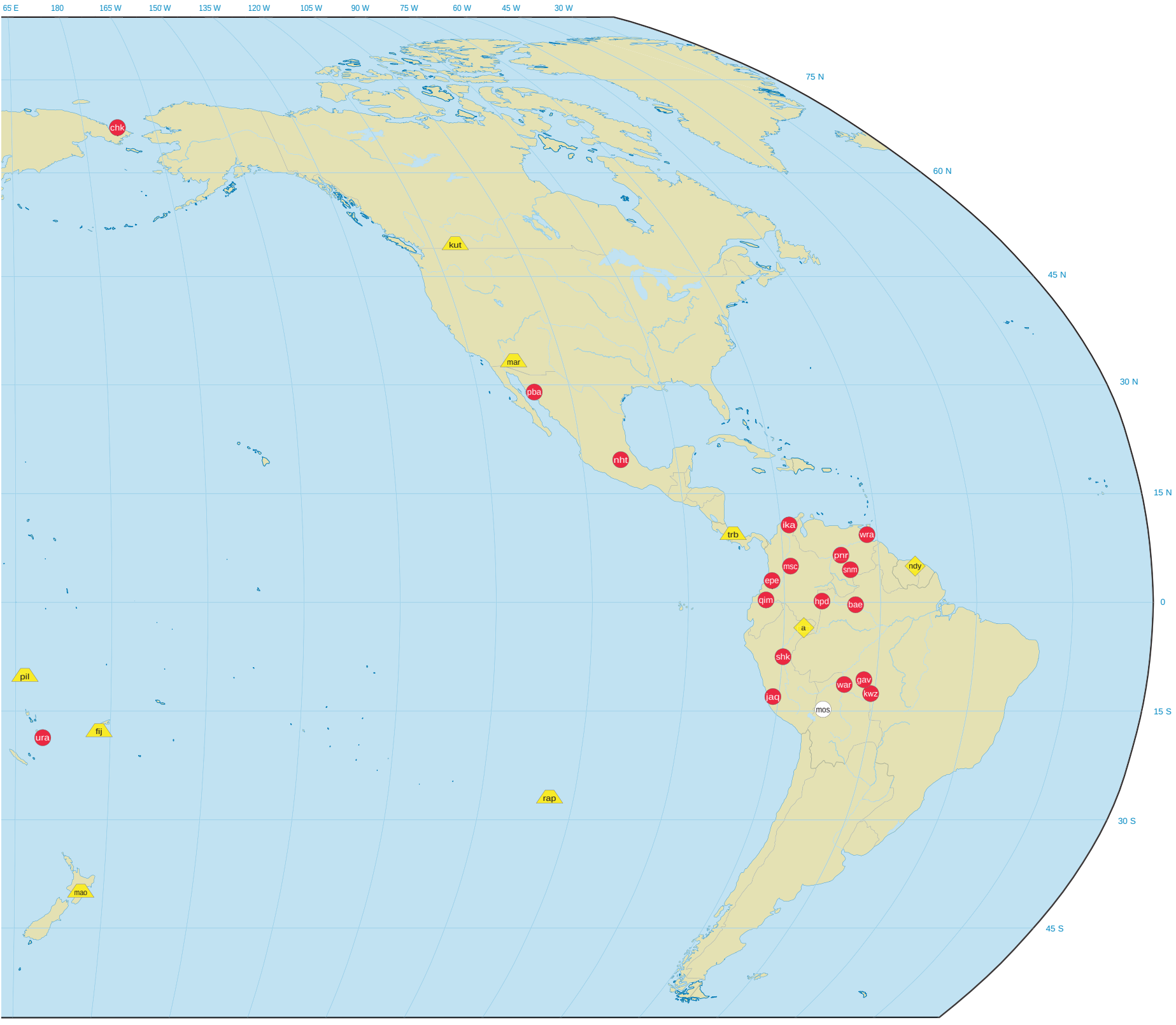


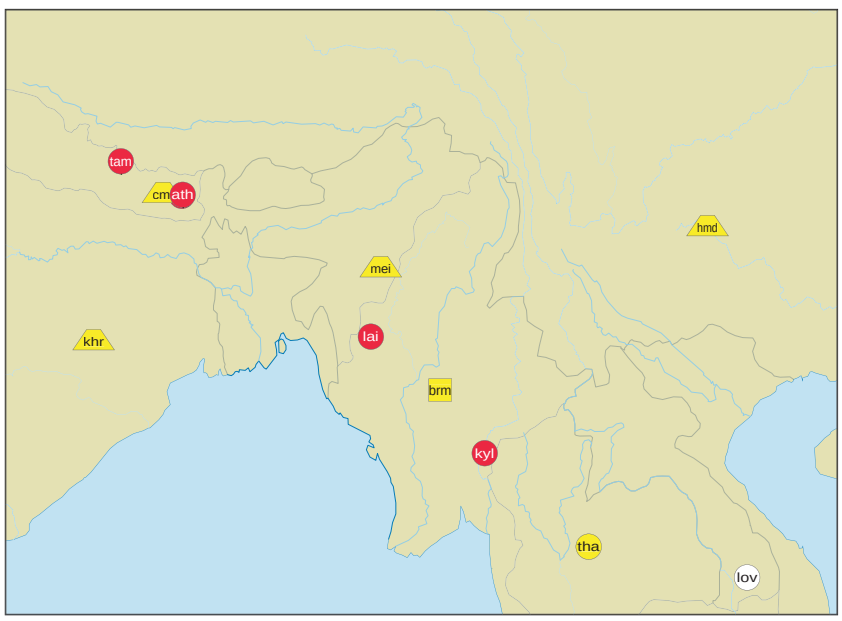


western Africa

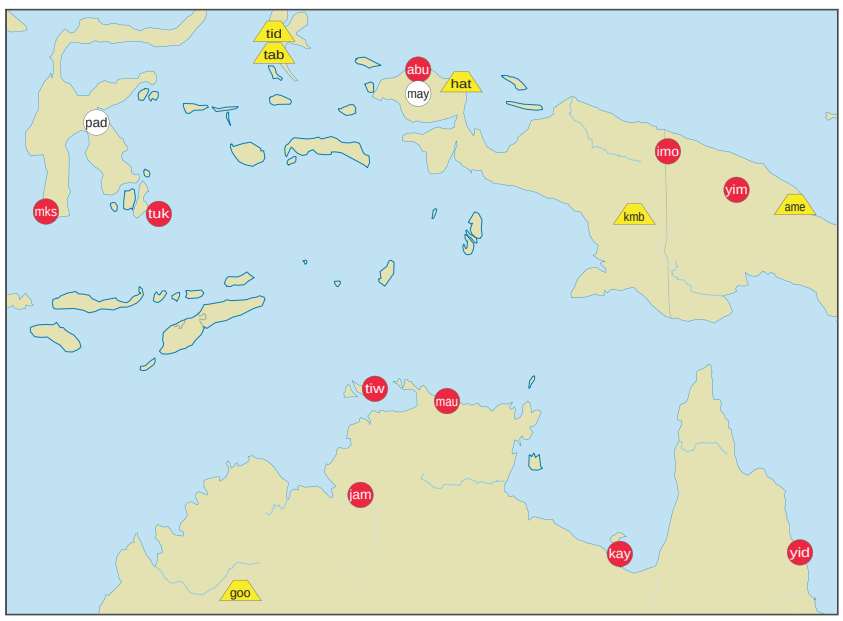


Caucasus and Middle East





Myanmar and surrounding areas



New Guinea and northern Australia

61 Adjectives without Nouns

DAVID GIL

1 Introduction

Adjectives may occur either as predicates, for example *This apple is **red***, or within noun phrases. When within noun phrases, they typically function as attributes to nouns, for example *I want the **red** apple*. However, in some cases, when the noun is either unimportant or is reconstructible from the discourse, it is absent from the construction, and, as a result, the adjective remains as the main lexical item within the noun phrase, denoting the understood object. An example of such a construction is the following:

(1) *I want the **red one**.*

In English, as suggested by the preceding example, adjectives without nouns generally occur in construction with a grammatical marker, the proform *one*. However, in some other languages, adjectives without nouns may occur with different kinds of construction markers, and in yet others they may occur without any such markers whatsoever. Moreover, in a small number of languages, adjectives without nouns do not occur at all. This map displays the distribution of adjective-without-noun constructions, and the kinds of markers such constructions contain.

2 Feature values

1. Adjective may not occur without noun	1
2. Adjective may occur without noun, and without marking	73
Adjective may occur without noun, obligatorily marked by	
3. prefix	7
4. suffix	13
5. preceding word	18
6. following word	7
7. mixed or other strategies	5
total	124

In languages of the first type, there are no adjectives without nouns; when occurring within a noun phrase, the adjective must always occur in construction with the noun that it qualifies. In the present language sample, there is but a single instance of such a language, Kayardild (Tangkic; Queensland, Australia; Evans 1995: 234).

The remaining six types contain languages that have adjective-without-noun constructions, providing a classification of their formal properties. In many languages, a bare adjective may occur without any additional construction markers, as a complete noun phrase denoting an understood object; such languages constitute the second type. Following is one such example:

(2) Hebrew (own knowledge)
Ani *roce* *et* *ha-adom.*
1SG want.PRES.M.SG ACC DEF-red
'I want the red one.'

Other examples of such languages include Lucazi (Fleisch 2001: 79–80), Basque (José Hualde p. c.), Burmese (Okell 1969: 84–7), Gaagudju (Harvey 1992: 318), Lillooet (Matthewson 1998: 217), and Bare (Aikhenvald 1996: 40).

The remaining five types contain languages in which adjectives without nouns cannot occur in bare form, but instead must occur with a particular marker that enables them to assume a more nominal function. Accordingly, such markers are commonly referred to as **nominalizers**. Of these five types, the first four provide a cross-cutting classification of the construction marker, depending on whether it is affixal (including clitics) or a separate word, and whether it precedes or follows the adjective that it marks.

The third type contains languages in which adjectives without nouns are marked with one or more prefixes. One such language is Semelai (Mon-Khmer; Malaysia), in which adjectives such as *tʰɔy* 'big' are preceded by the comparative marker *raʔ* and the relative marker *mə* (historically derived from the numeral 'one'), yielding forms such as *mə=raʔ-tʰɔy* 'big one', as in the following example:

(3) Semelai (Nicole Kruspe p. c.)
Jon *yɛ* *mə=raʔ-tʰɔy.*
give 1SG REL=COMPR-big
'Give me the big one.'

Other examples of such languages include Koyra Chiini Songhay (Heath 1999b), Hatam (Reesink 1999: 46), and Gavião (Moore 1984: 143–4).

The fourth type contains languages in which adjectives without nouns are marked with one or more suffixes. One such language is Kolyma Yukaghir (isolate; Siberia), in which adjectives such as *pojne* 'white' are marked with a participial suffix *-j* and then a nominalizing suffix *-ben*, resulting in forms such as *pojne-j-ben* 'white one', as in the following example:

(4) Kolyma Yukaghir (Elena Maslova p. c.)
Pojne-j-ben *lew-din* *erd'-ije.*
white-PTCP-NMLZ eat-INF want-1SG.INTR
'I want the white one.'

Other examples of such languages include Lezgian (Haspelmath 1993: 312–14), Imonda (Seiler 1985: 59, 185–93), and Yagua (Doris Payne p. c.).

The fifth type contains languages in which adjectives without nouns are marked with one or more preceding words. One such language is Iraqw (Cushitic, Afro-Asiatic; Tanzania), in which adjectives such as *ùr* 'big' are preceded by a dummy word in the construct state bearing the number and gender of the understood noun, yielding constructions such as (for a singular feminine object) *ar ùr* 'big one', as in the following example:

(5) Iraqw (Maarten Mous p. c.)
Ar *ùr* *a* *hlàa.*
CONSTR.F.SG big OBJ.FOC want.1SG
'I want the big one.'

Other examples of such languages include Albanian (Oliver Iggesen p. c.), Vietnamese (own fieldwork), and Maybrat (Dol 1999: 147, 307).

The sixth type contains languages in which adjectives without nouns are marked with one or more following words. One such language, illustrated in (1) above, is English; another is Mandarin, in which adjectives such as *hóng* 'red' are followed by an associative marker *de*, resulting in constructions such as *hóng de* 'red one', as in the following example:

(6) Mandarin (Violet Poo, Geraint Wong p. c.)
Wǒ *yào* *hóng* *de.*
1SG want red ASSOC
'I want the red one.'

Other examples of such languages include Hindi (Anvita Abbi p. c.), Korean (Hak-Soo Kim p. c.), and Ndyuka (George Huttar p. c.).

Finally, the seventh type contains languages in which adjectives without nouns are marked with combinations of the above types of strategies, or with yet other morphosyntactic devices. An example of the former kind is Eastern Kayah Li (Sino-Tibetan; Myanmar and Thailand), in which adjectives such as *bɛ̄* 'yellow' are marked with the nominalizing prefix *ʔa-* and also a following word *təplɔ* consisting of *tə-* 'one' plus classifier *plɔ*, yielding constructions such as *ʔabɛ̄ təplɔ* 'yellow one', as in the following sentence:

(7) Eastern Kayah Li (David Solnit p. c.)

V̄ɛ̄	*síjɯ*	*ʔa-bɛ̄*	*tə-plɔ.*
1SG	want	NMLZ-yellow	one-CLF

'I want the yellow one.'

An example of the latter kind is Epena Pedee (Choco; Colombia), in which the adjective loses its final vowel and undergoes a stress shift from final to penultimate, as in *pʰaimáa* 'black', *pʰáima* 'black one' (Harms 1994: 24).

3 Geographical distribution

The most salient geographical fact is the widespread distribution, in every major region, of the most common type, in which adjectives without nouns may occur in bare form, without any additional construction marker. In many parts of the world this type occurs intermingled with other types; however, in Australia it occurs almost exclusively.

Some of the other types occur with insufficient frequency to enable clear geographical patterns to be discerned. One clear pattern, though, is the widespread occurrence of the fifth and sixth types, involving periphrastic marking, across the eastern part of Asia. Moreover, within this region, there is a clear split, with following words concentrated in northeastern Asia and preceding words in South-East Asia. A possible additional cluster of the preceding-word strategy may be observed in parts of eastern Africa.

4 Theoretical issues

Adjectives without nouns represent a type of construction that has so far attracted relatively little attention in the general linguistic literature.

One set of questions involves the search for typological correlates of the various language types displayed in this map. Some correlations are obvious. For example, among the languages that do not allow adjectives without nouns to occur in bare form, the morpho-syntactic strategy to mark such nouns will be consistent with the overall morphological typology of the language: isolating, prefixing, or suffixing.

A more substantive proposed correlation suggests that languages will allow adjectives without nouns to occur in bare form if and only if adjectives are the target of morphological agreement controlled by the noun. While this may hold as a statistical tendency, counter-examples to this correlation can be adduced in both directions (Gil 1994d). Thus, for example, Hunzib is a language with adjectival inflection in which adjectives without nouns are obligatorily suffixed (van den Berg 1995: 57), while Minangkabau is a language without adjectival inflection in which adjectives without nouns may occur in bare form (own knowledge).

Another possible correlation suggests that languages will allow adjectives without nouns to occur in bare form if and only if adjectives are themselves noun-like in their grammatical behaviour. To be in a position to test such a hypothesis, it is necessary to identify specific characteristics of noun-like adjectives, distinguishing them from other kinds of adjectives, such as verb-like ones.

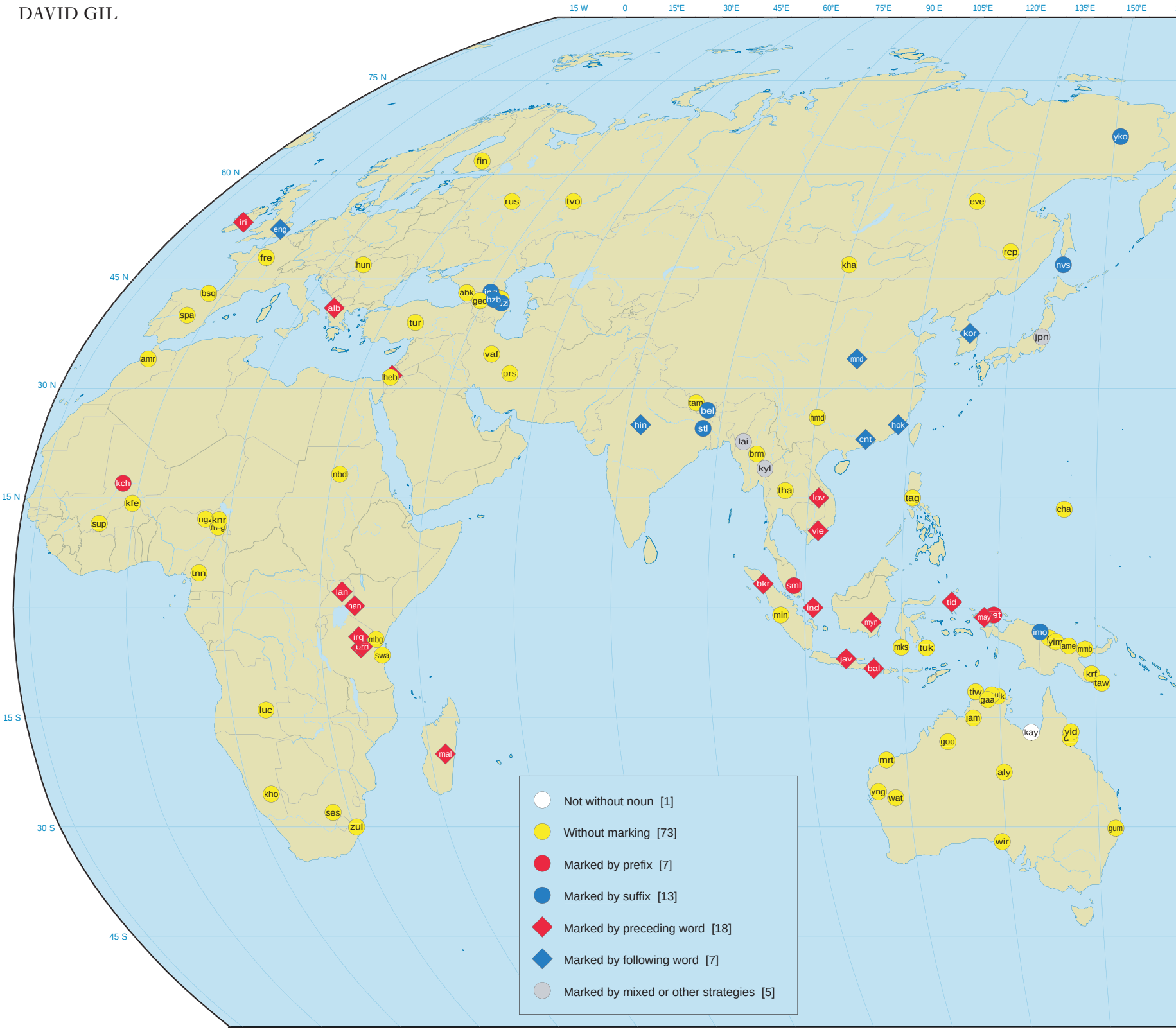
A second array of questions relates to the proper analysis of the adjective-without-noun construction. One issue, alluded to in the preceding paragraph, is the category membership of the adjective: is it really an adjective, or is it a noun, or perhaps a member of some other, less differentiated part of speech? In this chapter, we have been using the terms *noun* and *adjective* in the traditional manner, as labels for semantic categories rather than syntactic ones; however, such terminology sometimes has the unfortunate effect of obscuring important distinctions in the grammatical behaviour, across different languages, of words and phrases with similar meanings.

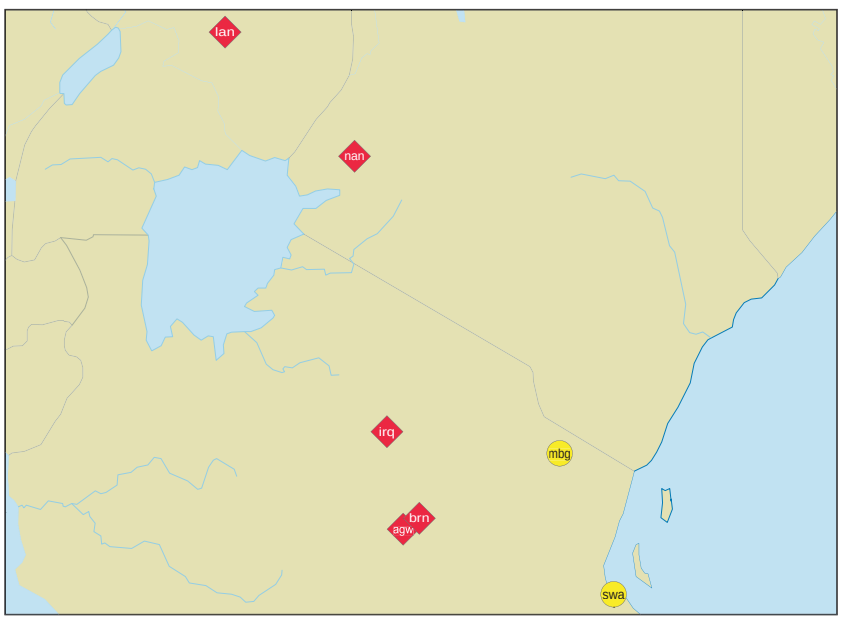
Another issue is that of headedness: whereas most linguists would agree that in a construction such as *red apple*, *apple* is the head and *red* its modifier, it is less obvious what the head is in a construction such as *red one* in (1). Since the grammatical marker *one* has many nominal properties, one might wish to argue that it, too, is the head of the construction. But what of the parallel Mandarin construction *hóng de* in (6)? Although in the present map, English *one* and Mandarin *de* are characterized similarly, as following words, they are quite clearly birds of a different feather. In particular, the Mandarin associative marker *de* has fewer nominal properties than English *one*; its functions are more clearly grammatical in nature. Accordingly, in Mandarin, the adjective would appear to have a stronger claim to be the head of the construction.

Thus, like all typologies, the one presented here focuses on certain features to the exclusion of others; it is just a preliminary step towards a better understanding of the adjective-without-noun construction.

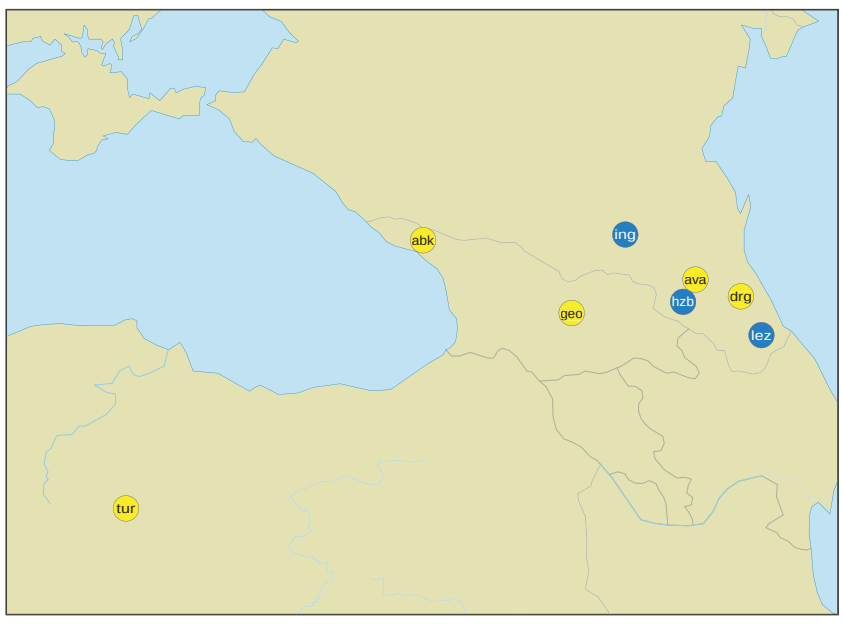
61 Adjectives without Nouns

DAVID GIL

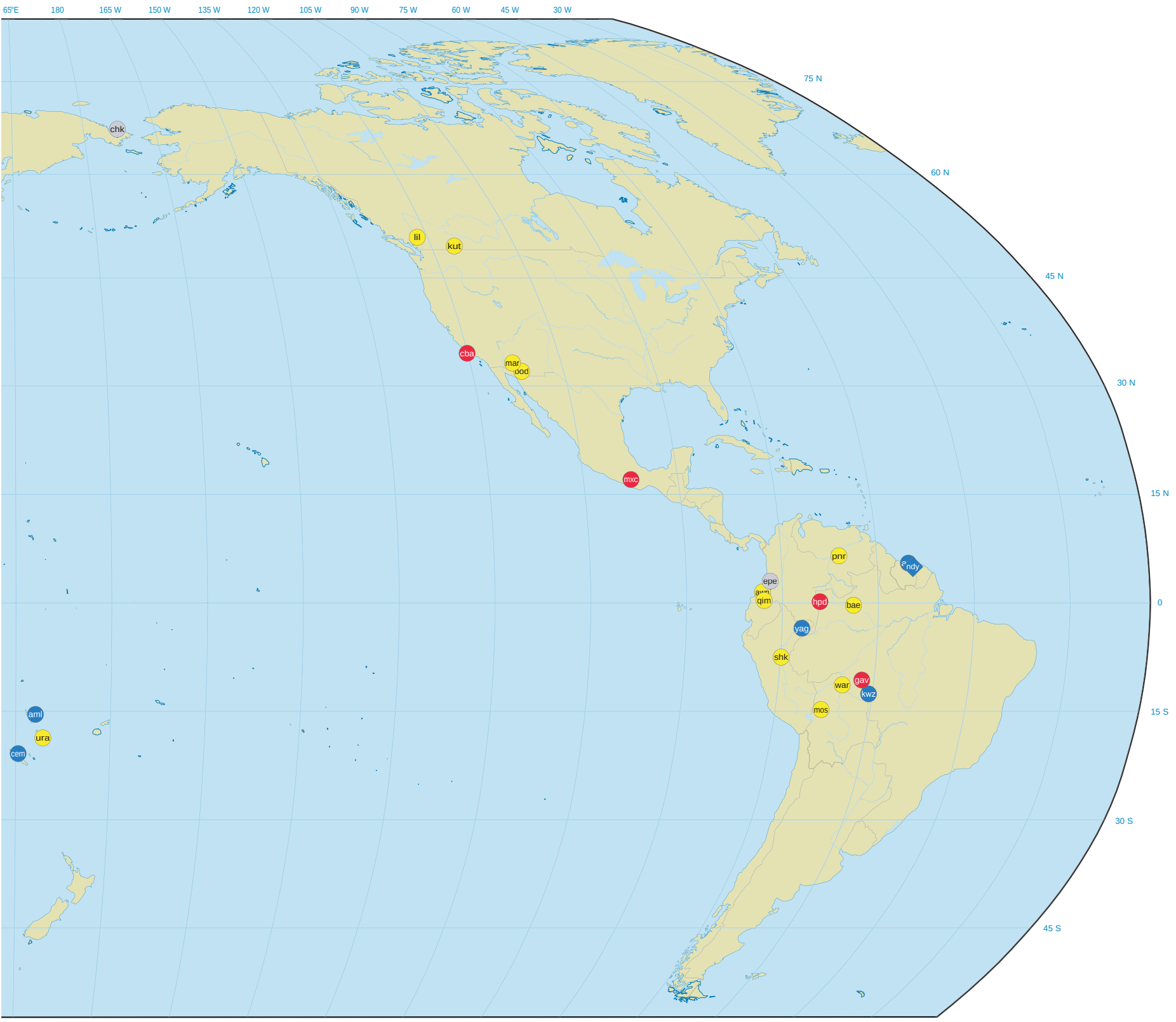


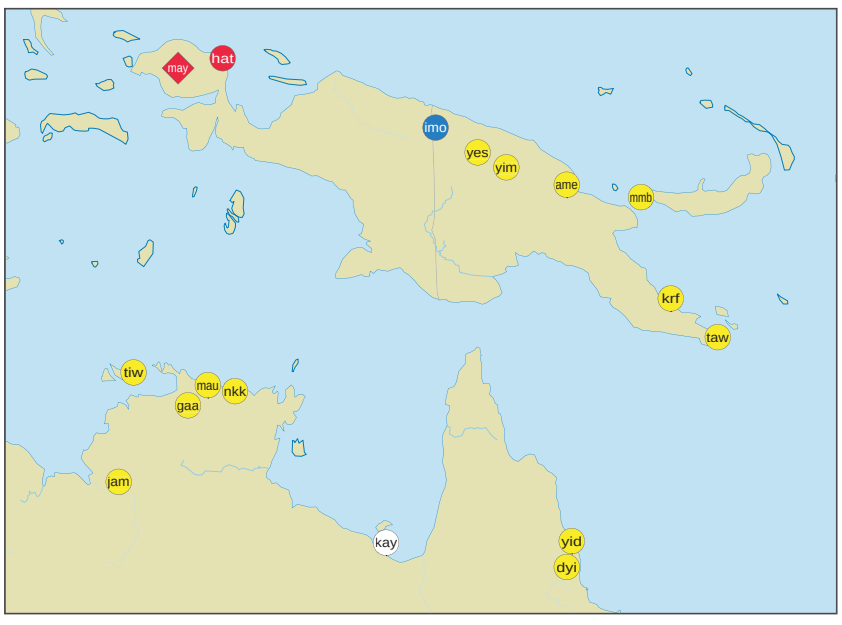


eastern Africa

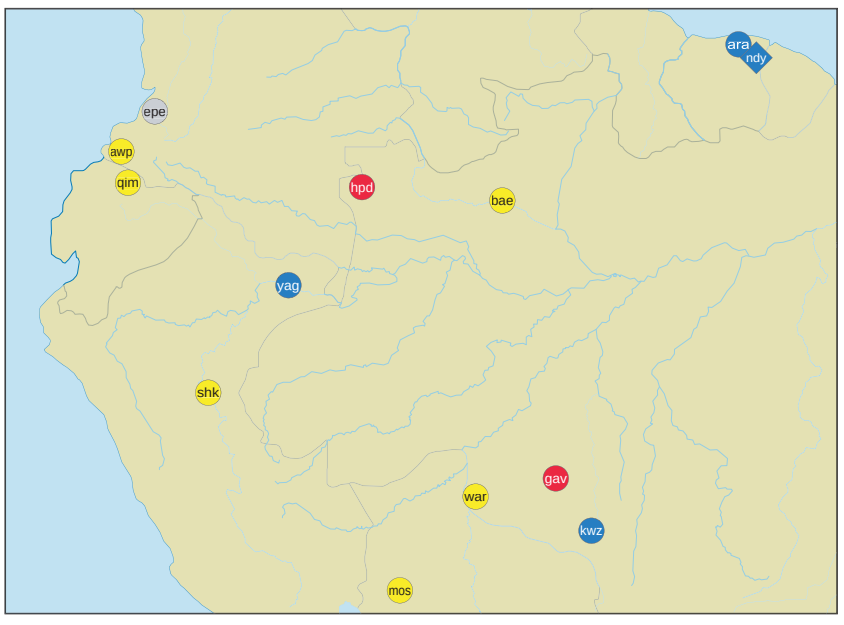


Caucasus and adjacent areas





New Guinea and northern Australia



northwestern South America

62 Action Nominal Constructions

MARIA KOPTJEVSKAJA-TAMM

1 Introduction

Map 62 shows the distribution across the world's languages of constructions corresponding to the English *John's running* and *the enemy's destruction of the city*. These are action nominal constructions (ANCs), i.e. constructions which have an **action nominal** (AN, *running*, *destruction*) as their head and contain a reference to the participants in the situation designated by that AN. ANs themselves either are nouns or at least occur in typical nominal positions and show inflectional properties and/or combinability with adpositions typical of nouns. They are, however, in some reasonably productive way derived from verbs, either derivationally or inflectionally, and refer to events and/or facts, i.e. not just to actions, as the name might imply. The issue at stake is whether words like *running* and *destruction* exist in a language at all and if so, how they combine with their arguments—such as *John* (the S argument), *the enemy* (the A argument), and *the city* (the P argument). (See also Chapters 98–99 for the notions S, A, and P.)

In the examples above, for instance, the ANCs are very similar to possessive noun phrases in that the S and the A appear as *'s*-possessors (cf. *Peter's house*) and the P appears as *of*-possessor (cf. *the house of my parents*). The whole ANC *the enemy's destruction of the city* looks very similar to the NP *the enemy's map of the city*, headed by an ordinary underived noun. On the other hand, these two ANCs differ significantly from the corresponding **finite clauses** *John runs* and *The enemy destroyed the city*, in which neither S nor A nor P needs any marker for signalling their syntactic functions. Such similarities between ANCs and possessive NPs, as opposed to the corresponding clauses, are by no means universal, and even in English there are other types of ANCs, such as *Peter's singing the Marseillaise* and *the destruction of the city by the enemy*, where the argument marking is to a certain extent analogous to that in finite clauses.

Comparing the internal structure of ANCs with that of NPs and finite clauses has been a hot issue in the last three to four decades, especially in the formalist tradition, and the literature is too vast even to be listed here. The main cross-linguistic studies of ANCs are Comrie (1976a), Comrie and Thompson (1985), and Koptjevskaja-Tamm (1993, 2002); the latter two works provide numerous cross-linguistic generalizations relating to ANCs.

There is a long tradition of distinguishing between **derivational action nominals** (like *conquest*, *refusal*, and *arrival*) and **inflectional action nominals**, like English gerunds (Chomsky 1970). This distinction will not be made here because it is precisely action nominals that in many languages pose serious problems for a clear-cut distinction between derivational and inflectional forms, in that the various criteria suggested for distinguishing inflection and derivation clash when applied to them (Comrie 1976a, Koptjevskaja-Tamm 1993: 263–6, and Haspelmath 1996). There is on the whole relatively little consensus concerning the terminology to be used for what are called **action nominals** and **action nominal constructions** in the present chapter. Thus, "verbal nouns", "gerunds", "nominalizations", "masdars", and "infinitives" often, though not necessarily, refer to ANs (Koptjevskaja-Tamm 1993: 22–43). Grimshaw (1990) uses the term "complex event nouns".

ANCs contrast with **clausal nominalizations**, which are not considered in this chapter. In ANCs, by definition, it is the head itself that has been turned into a noun, or nominalized, and thus represents a kind of lexical nominalization. In clausal nominalizations, on the other hand, it is the entire finite clause that gets nominal inflectional features, so that its head cannot be considered a noun (Comrie and Thompson 1985: 392), as in (1) from Imbabura Quechua.

(1) Imbabura Quechua (Cole 1982: 34)

Pedro	*ya-n*	*[ñuka*	*Agatu-pi*	*kawsa-ni]-ta.*
Pedro	think-3	I	Agato-in	live-1-ACC

'Pedro thinks that I live in Agato.'

More specifically, the verb in such nominalized clauses normally retains all its verbal characteristics (such as person and number inflection) and does not occur with nominal derivational affixes, although the nominalized clause as a whole may to some extent be treated like an NP. Clauses may be nominalized in different ways, for instance by taking an article or by attaching nominal inflectional suffixes to finite verbs, such as the accusative case in (1). In such cases, the verb inflects in the same way as a nonderived nominal; but in language-specific terms there may still be good arguments against considering it an action nominal—e.g. the "inflection of the verb" is actually clause-level inflection, which gets localized on the verb as the clausal head. In a few cases, however, the distinction between ANCs and clausal nominalizations is difficult to draw.

Most often ANCs serve as complements to predicates and refer to propositions, facts, or events (to use the semantic taxonomy launched by Vendler (1967)). Verbs such as *to assert* and *to believe* combine with ANCs referring to propositions, the verbs *to know* and *to regret* take ANCs referring to facts, while the verbs *to hear* and *to continue* take ANCs referring to events. The exact range of the meanings and functions typical for the ANCs, as well as their frequency in actual discourse, varies considerably across languages and can only be understood fully in the context of the overall system of clausal complementation in the given language. Accordingly, finding examples of ANCs may be more or less easy in different languages.

2 ANCs versus clauses versus noun phrases

ANCs show considerable variation in their internal structure, i.e. in the way ANs combine with their arguments. And there is a clear reason for variation here, both within a language and cross-linguistically: ANs are intermediate between typical verbs and nouns, and ANCs are intermediate between clauses and NPs.

Thus, verbs typically refer to events which involve various participants, they normally take arguments and function as predicates in clauses. Clauses can typically occur on their own, assert the occurrence (or nonoccurrence) of events, and locate them temporally. Nouns typically refer to persons, objects, places, and other more or less concrete entities. They function as heads in noun phrases, which are typically used as arguments to predicates (e.g. as subjects and objects). Nouns do not normally take arguments (with the possible exception of kin terms and other relational nouns), but often take various attributes. Possessive NPs, like *Peter's house*, cover numerous relations and constitute, in a sense, an unmarked, or default, option for combining a noun with one dependent nominal at a time.

Now, ANCs are semantically very close to finite clauses: thus, both *John's running* and *John runs* can describe one and the same event involving the same participant, and the same goes for *the enemy's destruction of the city* and *the enemy destroyed the city*. However, ANCs, as opposed to normal finite clauses, merely name the event and occur in functions typical of NPs, for instance as arguments to predicates (as in *The enemy's destruction of the city led to a complete collapse of the country's economy*). In these functions ANs themselves inflect or combine with adpositions in the same way as normal nouns do, while they may (and typically do) lack verbal characteristics such as mood and tense markers. There is, however, often a need to refer to the participants in the situation described

by an AN, i.e. to combine the AN with one or more arguments—a feature that clearly distinguishes them from normal nouns, which typically lack arguments. This closeness of ANs both to verbs and to nouns manifests itself in the structure of ANCs: in some languages ANCs look more or less like other NPs, while in others they show clear similarities to finite clauses.

Thus, cross-linguistically ANCs may be classified in accordance with their sentence-likeness and/or NP-likeness, i.e. with the degree to which the relations between an AN and its arguments are signalled in the same way as verb-argument relations in finite clauses ("sentential marking") or as head-dependent relations in NPs. The typology of ANCs assumed here is based on the marking of the arguments themselves (by means of case marking and adpositions) as a primary parameter, in a few cases supplemented by head marking as an additional parameter (see Koptjevskaja-Tamm 1993: 58–60 for extensive discussion). The main cross-linguistic generalization here is as follows:

ANC Universal: ANCs as parasites
In all languages where ANCs exist, they use marking (both dependent marking and head marking) which is also used in either finite clauses or possessive NPs.

In other words, ANCs, in a sense, do not exist as autonomous, independent constructions, but always "borrow" syntactic means from finite clauses and/or NPs.

3 Defining the values

The following feature values are distinguished in Map 62:

●	1. Sentential: dependent marking of the finite clause is retained for S, A, and P	25
●	2. Possessive-Accusative: S/A treated as possessors, P retains sentential marking	29
●	3. Ergative-Possessive: S/P treated as possessors, A treated differently	21
●	4. Double-Possessive: All major arguments treated as possessors	7
◇	5. Other: minor patterns	6
○	6. Mixed: several patterns in the same language	14
●	7. Not both A and P in the same construction	24
○	8. No action nominals	42
	total	168

All the values will be illustrated by transitive ANCs, i.e. those involving A and P. The first value (the **Sentential** ANC type) is exemplified by Godoberi (Daghestanian; eastern Caucasus). Example (2) shows a (di-)transitive ANC where the major syntactic functions are dependent marked according to the same ergative-absolutive pattern as would be involved in a corresponding finite clause: S and P are in the nominative, A is in the ergative, and the recipient is in the dative case.

(2) Godoberi (Kazenin 1996: 160)
aHmadi-di *maHamadi-li* *rec'i* *ik̄-ir*
Ahmad-ERG Mahamad-DAT bread.ABS give-AN
'Ahmad's giving bread to Mahamad'

Example (3) from Meadow Mari (Uralic; Russia) illustrates value two (the **Possessive-Accusative** type), whereby the A (and the S) are treated as possessors, while the P retains its sentential form. In (3), the A takes the genitive case and the AN attaches the corresponding possessive suffix, according to the general pattern of possessive NPs in the language, while the P takes the accusative case used for direct objects in finite clauses.

(3) Meadow Mari (Natal'ja Serdobol'skaja, p. c.)
möj-ön *pis'ma-m* *voz-öm-em*
I-GEN letter-ACC write-AN-1SG.POSS
'my writing of the letter'

Example (4) from Russian illustrates value three (the **Ergative-Possessive** type), whereby the P (and the S) are treated as possessors, here appearing in the genitive case, while the A is treated differently, as a more peripheral NP. Cross-linguistically the A will typically, though not universally, appear in the agent case of passive sentences, as in (4).

(4) Russian
ispoln-enij-e *sonat-y* *pianist-om*
perform-AN-NOM sonata-GEN pianist-INSTR
'the performance of the sonata by the pianist'

Value four (the **Double-Possessive** type), whereby all the S, A, and P in ANCs are treated as possessors, is illustrated by example (5) from Tukang Besi (Austronesian; Sulawesi, Indonesia).

(5) Tukang Besi (Mark Donohue, p. c.)
te *basa-'a* *nu* *La* *Petrus* *nu* *boku*
CORE read-AN POSS PROP.M Peter POSS book
'Peter's reading of the book'

Value five encompasses several minor patterns, each having a very restricted cross-linguistic distribution. Thus, in (6) from Chamorro (Austronesian; Micronesia), the A ('Juan') is treated as a possessor, while the P ('tuba') appears with the preposition *ni*, otherwise used for various kinds of obliques, such as instruments and passive agents.

(6) Chamorro (Topping 1973: 221, Sandra Chung, p. c.)
i *ginimen* *Juan* *ni* *tuba*
DEF.ART drink.AN Juan OBL tuba
'Juan's drinking of the tuba (= a traditional drink)'

Value six is assigned to languages that have several different types of ANCs. For instance, English has ANCs of the Possessive-Accusative type (*Peter's singing the Marseillaise*), of the Ergative-Possessive type (*the destruction of the city by the enemy*), and of the Double-Possessive type (*the enemy's destruction of the city*).

A number of languages have action nominals, but these can never combine with both A and P within one and the same construction—these languages are shown with value seven on the map. Sometimes there are strategies for expressing both A and P at the same time without, however, making both of them syntactically dependent on the AN. Thus, the English ANC *X's buying of the dress* may be rendered by constructions like "X's dress-buying" (where the P and the action nominal form a compound), or "the buying of the dress that X did" or "the buying that X bought the dress". The latter option is illustrated by (7) from Yoruba.

(7) Yoruba (Ayo Bamgbose, p. c.)
rírà *tí* *Olú* *ra* *aso*
buy.AN REL Olu buy dress
'Olu's buying of the dress'

Finally, some languages completely lack action nominals as a class—these have the value eight. Their closest correspondents to ANCs vary: examples are finite subordinate clauses in Chalcatongo Mixtec, clausal nominalizations in Slave and many other North American languages, and verbal nouns with defective paradigms in Gagauz (Turkic; eastern Balkans).

4 Geographical distribution

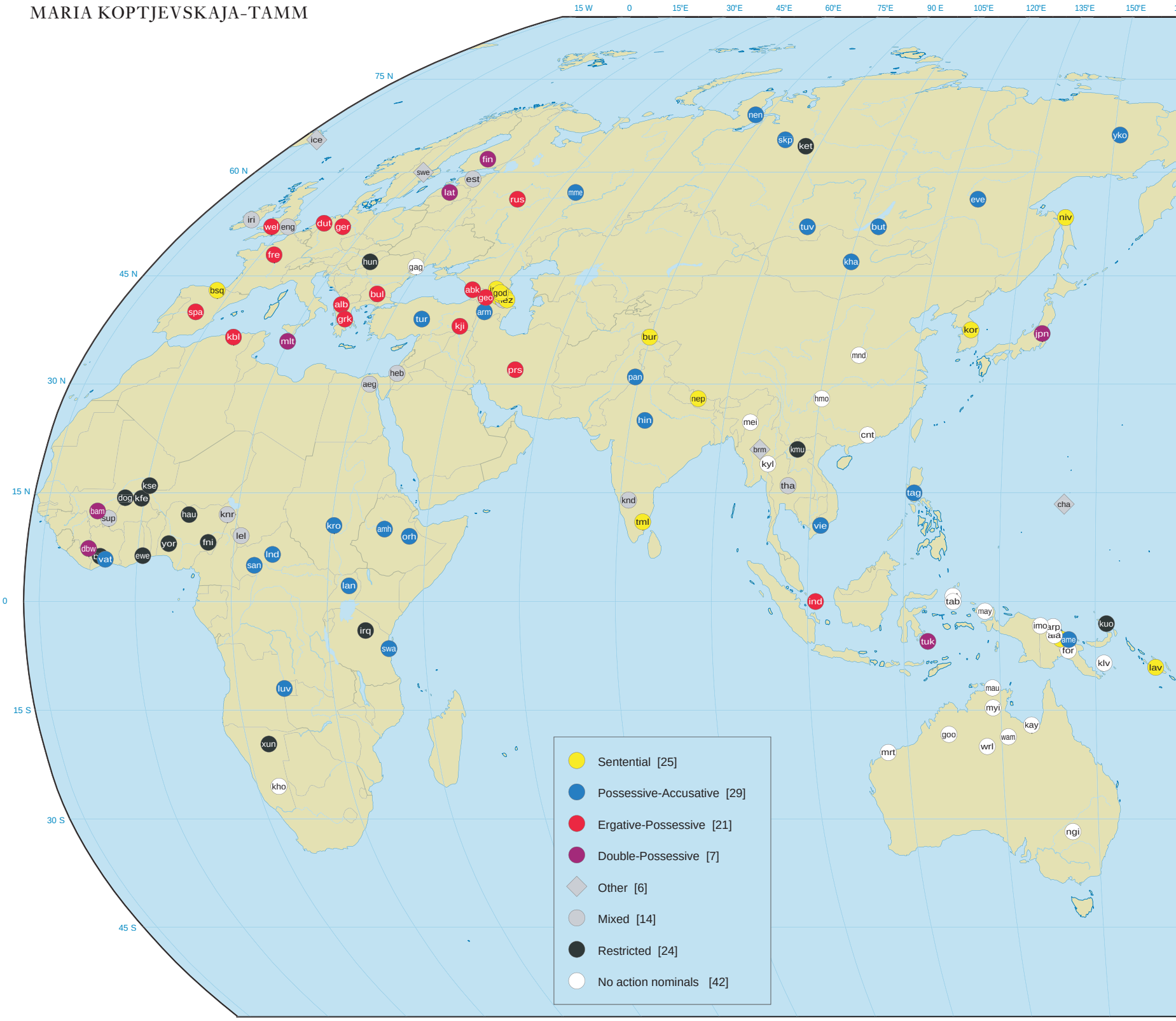
The most striking and clearest areal tendency concerns the existence of ANs themselves. Australian languages seem to lack ANs, while often having other deverbal derivatives. North America is another area where ANs, and in particular ANs combining with arguments, are a rare phenomenon. ANs are common in both Europe and Africa, but in a number of African languages they cannot combine with A and P at the same time.

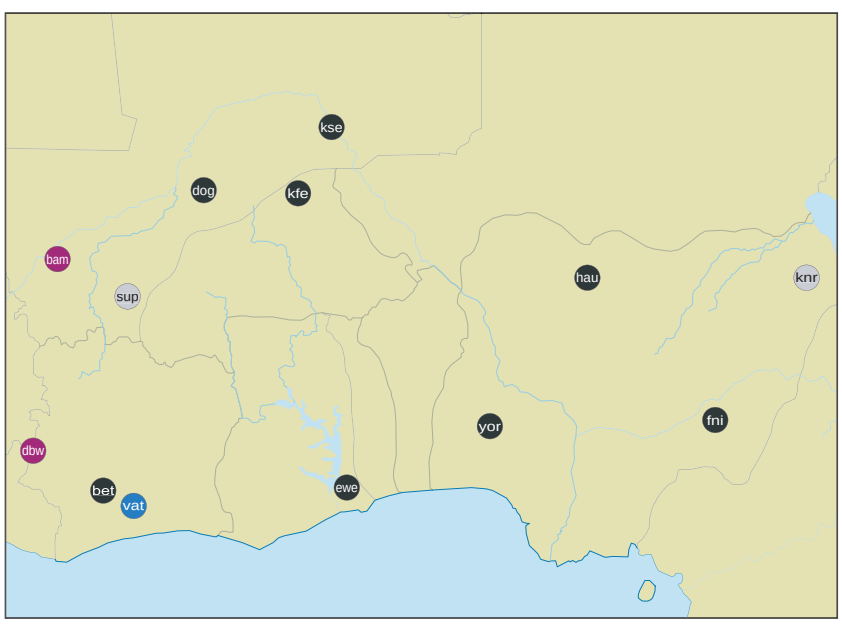
Acknowledgement

The work presented here was supported by grant (2002–1478) from the Swedish Research Council.

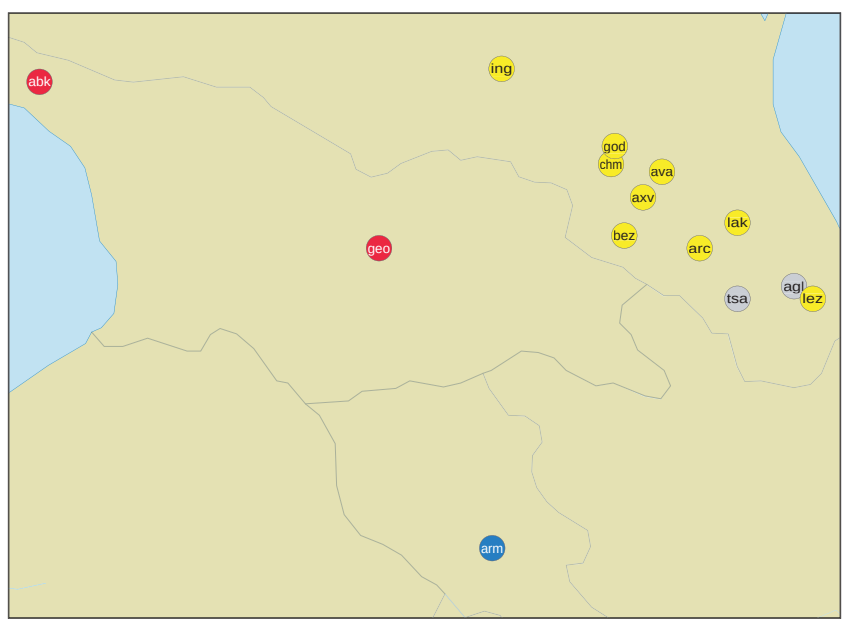
62 Action Nominal Constructions

MARIA KOPTJEVSKAJA-TAMM

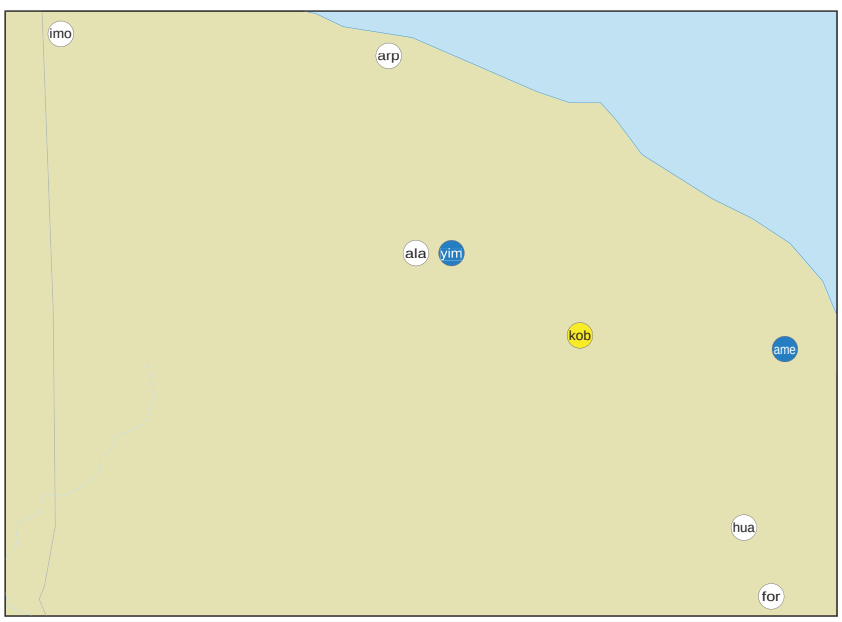




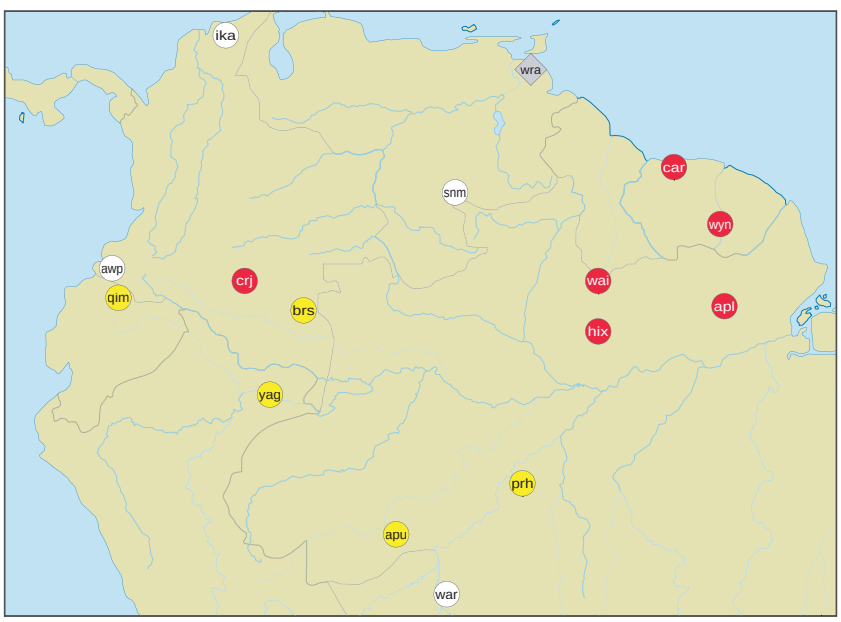
western Africa from Liberia to Cameroon



Caucasus



eastern New Guinea



northern South America

63 Noun Phrase Conjunction

LEON STASSEN

1 Defining the values

This map shows the distribution of the variation in the encoding of noun phrase conjunction. The basic distinction is between those languages which use a different marker for noun phrase conjunction and comitative phrases (so-called AND-languages; Stassen 2000) and those languages in which the markers for noun phrase conjunction and comitative phrases are the same (WITH-languages). A clear example of an AND-language is English. As the examples in (1) demonstrate, in this language the marker of noun phrase conjunction (*and*) is different from the marker of comitative phrases (*with*). In contrast, the example in (2) shows that Nkore-Kiga (Bantu; Uganda) is a WITH-language: the item *na* can be translated as either 'and' or 'with'. An alternative way of stating this is to say that WITH-languages like Nkore-Kiga lack the structural option of noun phrase conjunction.

(1) a. *John and Mary went to the movies.*
b. *John went to the movies with Mary.*

(2) Nkore-Kiga (Taylor 1985: 58)
n-ka-za-yo *na* *Mugasho*
1SG-REC.PST-go-there and/with Mugasho
'Mugasho and I went there./I went there with Mugasho.'

Accordingly, the following two values are shown on the map:

1.	AND-languages: 'and' and 'with' are not identical	131
2.	WITH-languages: 'and' and 'with' are identical	103
	total	234

2 Variation in noun phrase conjunction marking

Among the languages which have the option of structural noun phrase conjunction, it is possible to subcategorize conjoined NP structures on the basis of the linking device which they employ. First of all, there are languages with **juxtaposition** or **zero marking** of such structures. Traditionally, the term **asyndeton** is employed to refer to such constructions. Quite a few languages can be shown to have this encoding possibility as at least one of their options. However, closer inspection reveals that asyndetic NP conjunction is nonetheless a minor strategy. Obligatory use of this variant is very rare. This option can be documented only in a few languages, among which no significant areal cohesion can be defined. An example is Awtuw (Lower Sepik-Ramu; Papua New Guinea).

(3) Awtuw (Feldman 1986: 67)
Yowmen *Yawur* *du-k-puy-ey*
Yowmen Yawur DUR-IMPF-hit-IMPF
'Yowmen and Yawur are hitting (someone).'

Zero-marked NP conjunction appears to be absent from Africa and from at least the western part of Europe. In all other parts of the world it must have been an old encoding option. However, the general trend all over the world is that zero conjunction tends to be marginalized into specific functions or is replaced altogether by overt marking strategies (see Mithun 1988a).

By far the most prominent option in overt marking of noun phrase conjunction is the use of a **medial connective**, i.e. a conjunctional item which stands between the two noun phrases. This strategy is illustrated by the English item *and* (see 1a). Noun phrase conjunction via a medial connective can be encountered all over the globe. In some areas, notably sub-Saharan Africa, the strategy is a minor one, but there is hardly any major linguistic area in which it is totally lacking.

In comparison to medial connectives, other overt conjunction strategies are fairly rare. All the strategies at issue feature postposition or suffixation of the connective item or items. Among these options, **polysyndeton** (i.e. the use of two conjunction markers) appears to be the most popular. This strategy can be found in a number of unconnected linguistic areas such as the Caucasus, northeastern Africa, Australia and New Guinea, southern India, and northeastern Asia. In addition, we find isolated examples in the Americas, in West Africa, and in Myanmar. An example of the strategy is Tamil (Dravidian; southern India):

(4) Tamil (Asher 1982: 69)
akkaa-vum *tangkacciy-um*
elder.sister-and younger.sister-and
'elder sister and younger sister'

Monosyndetic postposing of conjunction markers (whether to the first or to the second conjunct) is low in frequency. Moreover, languages which present this option typically also allow a construction of the polysyndetic type, so that these monosyndetic constructions are best regarded as variants in which one of the markers in the polysyndetic construction is optionally deleted. Given this, it will be clear that the two monosyndetic variants occur in roughly the same areas as the polysyndetic construction. An example of the "first-NP" subtype is Beja (Cushitic; Sudan), while the "second-NP" subtype is illustrated by NP conjunction from Tubu (Nilo-Saharan; Chad and Niger).

(5) Beja (Reinisch 1893: 195)
a. *ani-wa* *baruk-wa*
1SG-and 2SG-and
'you and I'
b. *mek-wa* *laga*
donkey-and calf
'a donkey and a calf'

(6) Tubu (Lukas 1953: 166)
a. *túrku* *ye* *mɔlɔfúr* *ye*
jackal and hyena and
'the jackal and the hyena'
b. *wúdɛn* *arkɔ́* *ye*
antelope goat and
'the antelope and the goat'

In the map, variation in the encoding of noun phrase conjunction has not been taken into account.

3 Variation within the comitative strategy

In contrast to noun phrase conjunction, the encoding of comitative phrases is much more formally uniform across languages. In the overwhelming majority of languages, the comitative strategy manifests itself by way of an oblique marker 'with' on the noun phrase. Depending on the general morphosyntactic features of the language the comitative marker may take the form of an adposition or an affix.

Although "dependent marking" (Nichols 1986) is the predominant encoding option in comitative constructions, there are some languages which deviate from this pattern. A "head-marking" strategy, in which the comitative marker is incorporated into the predicate, is encountered in North-West Caucasian languages such as Ubykh.

(7) Ubykh (Dumézil 1931: 17)
go-u-ji-k'ä-qa
3SG-2SG-WITH-come-PERF
'He came with you.'

In some cases, this incorporation leads to transitivization of the predicate, so that the "comitative" NP takes on the syntactic status of a direct object. Assorted languages of the Americas present this option; the example is from Selknam (Chon; Patagonia, Argentina).

(8) Selknam (Tonelli 1926: 58)
Kokoš telken okel-enen.
Kokoš boys with-go
'Kokoš goes with the boys.'

The transitive patterning of comitatives is manifested even more explicitly in languages where the comitative NP is constructed as the direct object of a verb 'to accompany/to have/to hold/to take/to follow' in a serialization construction. Clearly, such cases will be found in areas where verb serialization is rampant anyway. Thus, West Africa, eastern Austronesian, and Sino-Tibetan are the places to look for this phenomenon. An example is Igbo (Igboid; Nigeria).

(9) Igbo (Welmers 1973: 369)
há sò anyị́ gá-á Ábá
they accompany.STAT us go.to-NARR Aba
'They went to Aba with us.'

Again, this variation in the encoding of comitative phrases is not reflected on the map.

4 Shift of WITH to AND

Although both for WITH-languages and for AND-languages clear instances can be found, the typological status of these two types is probably not equally well established. In particular, there is a notable discrepancy in the stability of the types. In general, AND-languages can be said to be stable diachronically and "pure" in their synchronic state: there is a sharp delineation between the two available strategies. On the other hand, "pure" instances of WITH-languages are relatively rare. For a considerable number of such languages, some process of "diachronic drift" of the comitative encoding can be attested. The general outcome of this process is to effectuate a shift from one-strategy to two-strategy status. These languages tend to introduce a secondary differentiation into the structural features of the comitative strategy, by changing one or more features of that strategy towards the features of the conjunctional strategy. In this way, the language acquires a two-strategy encoding of the domain, in which one of the strategies is still purely comitative, while the other is some sort of hybrid between the comitative and the conjunctional strategy.

The "shift" of the comitative strategy in WITH-languages is a gradual process. Moreover, the structural features which are changed in the course of this process are different for different groups of WITH-languages. Among the possible "paths" involved in the shift, we encounter cases in which:

(i) The comitative NP can be moved from its canonical position to a position adjacent to the other conjunct, so that the two conjuncts now form a constituent. An example is Babungo (Bantoid; Cameroon).

(10) Babungo (Schaub 1985: 87)

a. *Làmbí gə̀ táa yìwìŋ ghɔ́ Ndùlá*
Lambi go.PFV to market with Ndula
'Lambi went to the market with Ndula.'

b. *Làmbí ghɔ́ Ndùlá gə̀ táa yìwìŋ*
Lambi and/with Ndula go.PFV to market
'Lambi and Ndula went to the market.'

(ii) There is differentiation in person/number agreement (mainly between singular and dual/plural). An example is Tolai (Oceanic; New Britain, Papua New Guinea).

(11) Tolai (Mosel 1984: 176)

a. *nam ra tutana i ga rovoi ma ra pap*
DEM ART man 3SG PFV hunt with ART dog
'That man went hunting with his dog.'

b. *Terengai dir rovoi ma ra pap*
Terengai 3DU hunt and/with ART dog
'Terengai and his dog went hunting.'

(iii) There is doubling of the comitative marker when used as a conjunction, thereby overtly indicating the equality in rank of the two noun phrases. A case in point is Japanese.

(12) Japanese (Hinds 1986: 97, 94)

a. *Taroo wa Akiko to Nara e ikimashita*
Taroo TOP Akiko with Nara to go.PST
'Taroo went to Nara with Akiko.'

b. *Taroo to Akiko to wa Nara e ikimashita*
Taroo with Akiko with TOP Nara to go.PST
'Taroo and Akiko went to Nara.'

Since, however, in all these "shift"-languages the markers for coordination and comitatives remain identical, such languages have been coded as WITH-languages for the purposes of this map.

5 Geographical distribution

As the map demonstrates, the distinction between the two language types shows clear areal features. Concentrations of WITH-languages can be found in two large linguistic areas. First, we find such languages in East and South-East Asia and the islands of Indonesia, Melanesia, and Polynesia. Secondly, most of the languages of Africa in and below the Sahara (with the notable exception of Khoisan) have WITH-status.

Opposed to this, AND-languages appear to be the rule in a macro-area which might roughly be called "Eurasia". This area comprises all the languages of Europe, Central Asia, and Siberia; in the east, it also includes some, though not all, of the Austronesian languages of the Philippines. To the south-east, the area stretches as far as India. Finally, it includes the Middle East and northern Africa. A second macro-area of AND-languages is formed by (most of) the languages of Australia and the central highlands of New Guinea.

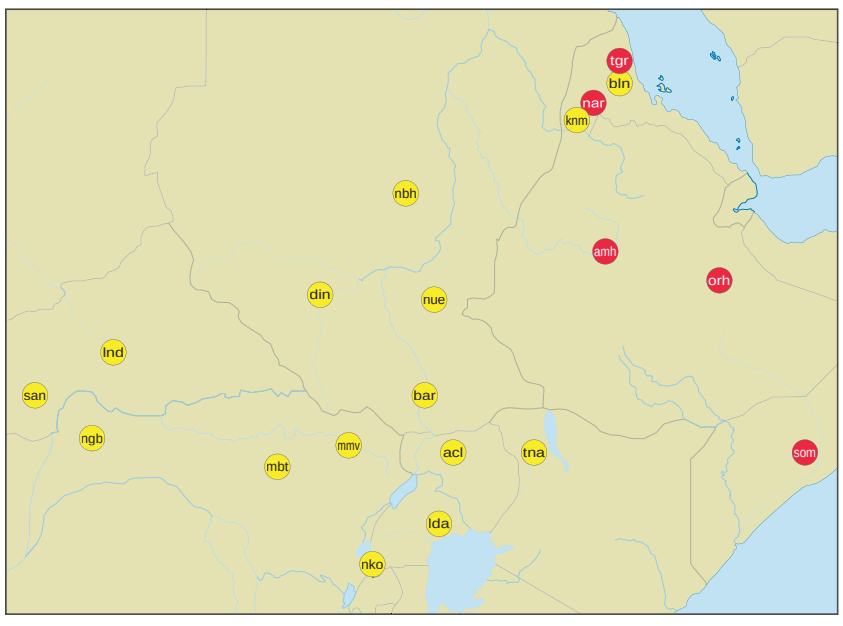
Outside of these two main AND-areas, AND-encoding can be found in several other places, in particular in the Americas, but the distribution of AND-languages and WITH-languages seems to be rather haphazard here. North America seems to have a predominance of AND-status, whereas the languages of the southern part of Central America and the eastern part of South America appear to tend towards WITH-encoding.

63 Noun Phrase Conjunction

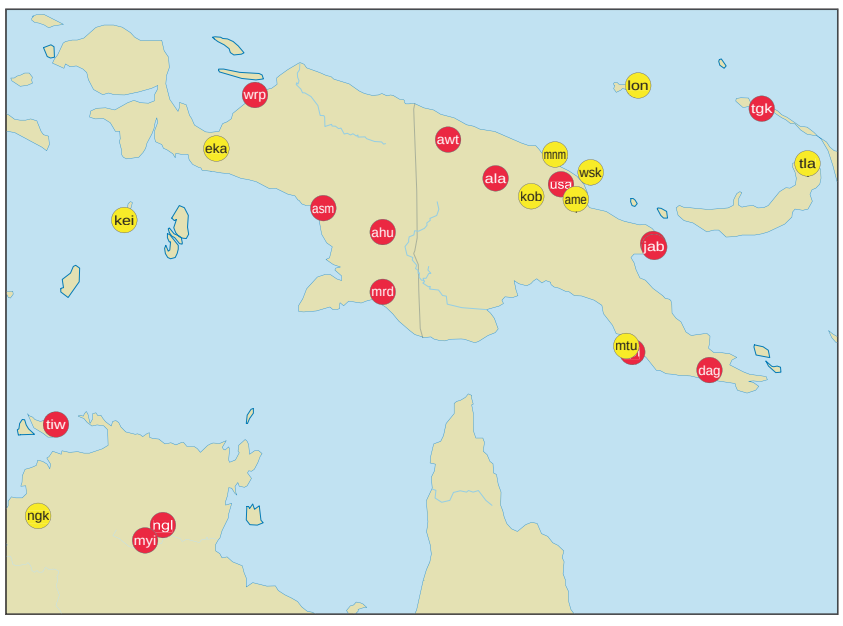
LEON STASSEN

‘And different from ‘with’ [131]

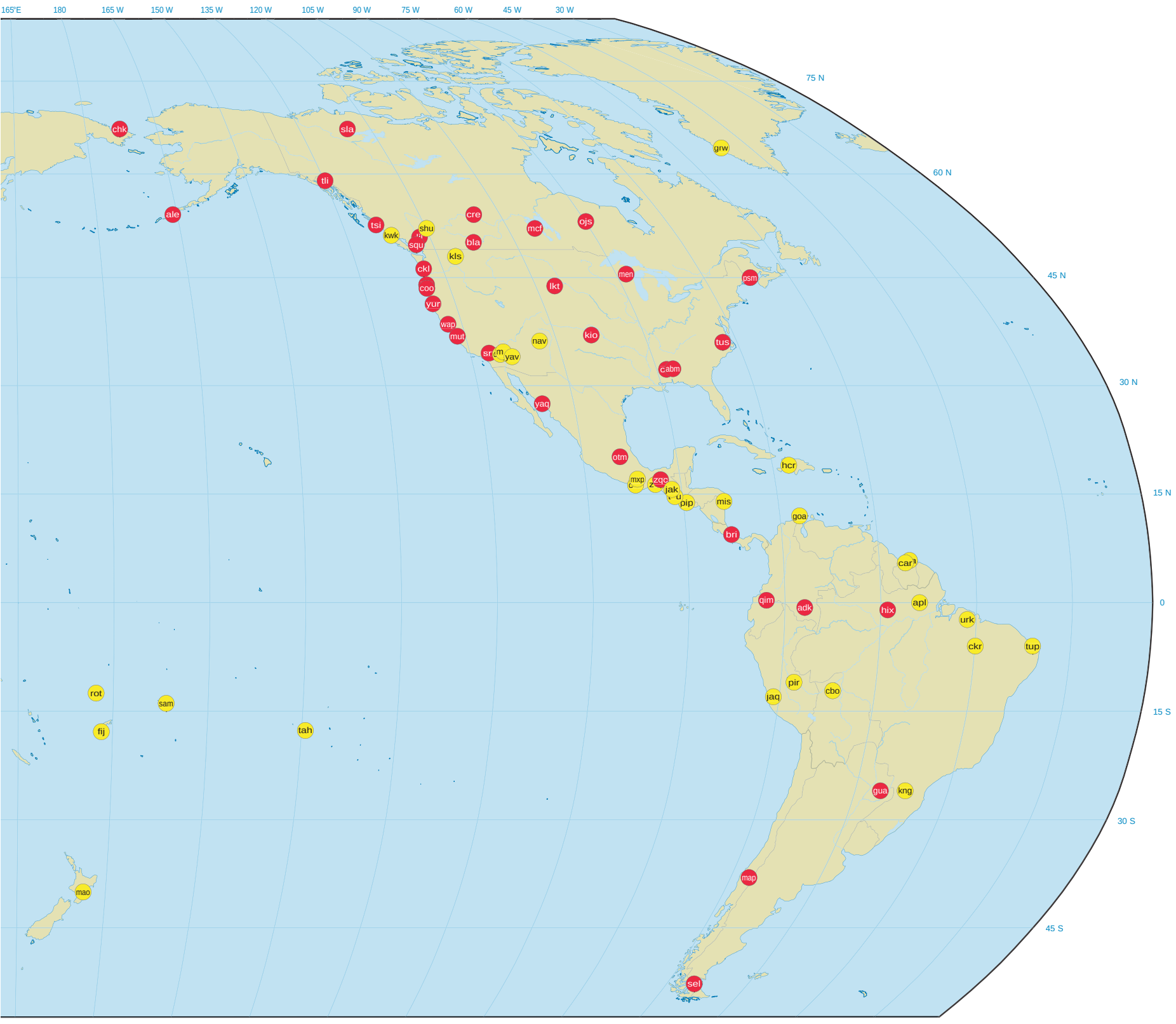
‘And identical to ‘with’ [103]

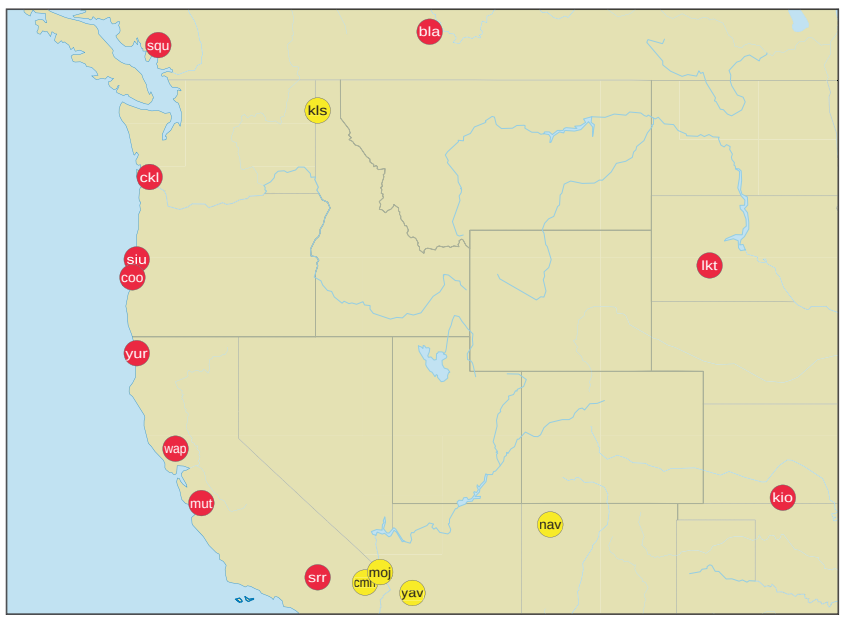


eastern Africa

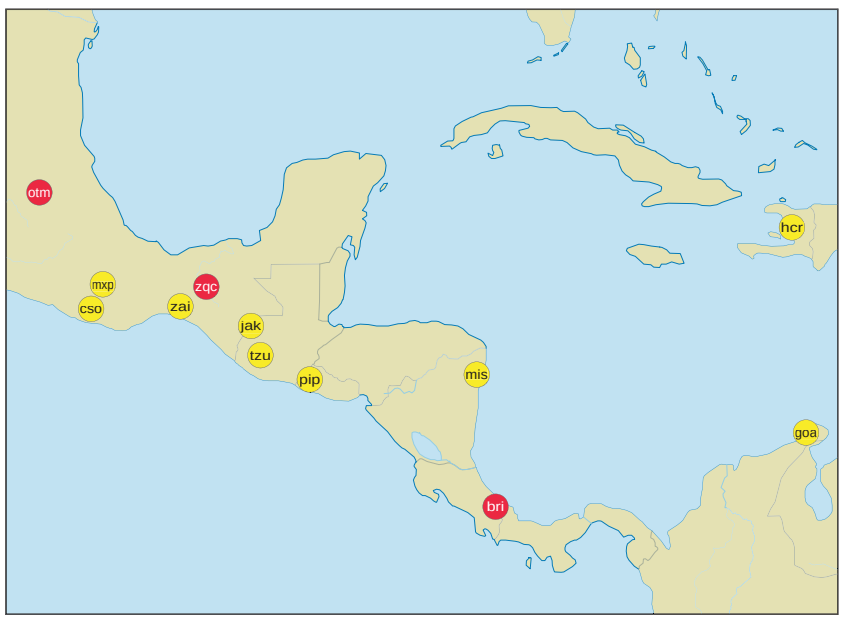


New Guinea and northern Australia





western United States



southern Mexico and Central America

64 Nominal and Verbal Conjunction

MARTIN HASPELMATH

1 Defining the values

In the European languages, the same marker is used both for **conjunction of noun phrases** and for **conjunction of verb phrases and clauses**, as illustrated in (1) for Hungarian (and for English in the translation). Thus, Hungarian *és* and English *and* can conjoin both nominal and verbal/clausal constituents.

(1) Hungarian (Kenesei et al. 1998: 102, 106)

a. *Anna* ***és*** *Péter*
Anna and Péter

b. *Péter olvas* ***és*** *tanul.*
Péter reads and studies

c. *Péter olvas* ***és*** *Anna tanul.*
Péter reads and Anna studies

But in many non-European languages, different conjunction markers are used for different constituent types. An example comes from Dagbani (Gur, Niger-Congo; Ghana), where *mini* conjoins noun phrases (cf. 2a) and *ka* conjoins verb phrases (cf. 2b) and clauses (cf. 2c).

(2) Dagbani (Olawsky 1999: 44, 51)

a. *doo ŋɔ* ***mini*** *m ba*
man this and my father
'this man and my father'

b. *O biɛ* ***ka*** *kɔɣisi.*
he be.bad and be.thin
'He is bad and thin.'

c. *Gbuɣima ŋubiri nimdi*
lion.PL chew.IMPF meat
ka *jansi diri kɔdu.*
and monkey.PL eat.IMPF banana
'Lions eat meat and monkeys eat bananas.'

Since clauses generally contain a verb phrase, I simply use the term *verbal conjunction* for both conjunction of verb phrases and conjunction of clauses.

The main contrast is thus between **identity** and **differentiation** of nominal and verbal conjunction. However, it seems useful to single out one further type: languages in which there is no overt marker either for nominal or for verbal conjunction, so that conjunction is expressed exclusively by **juxtaposition**. Strictly speaking, this is a subtype of the first type, but it is treated separately here because of its interesting geographical distribution. Example (3) is from Nhanda (Kardu, Pama-Nyungan; Western Australia).

(3) Nhanda (Blevins 2001: 134)

a. *acijadi-wana mirla-wana*
clothes-1PL rug-1PL
'our clothes and our rugs'

b. *Nyini yatka-da mutha=ngga ngayi urndu=galu.*
you go-IMPF nose=LOC I back=PATH
'You go in the front and I'll go behind.'

However, when just one of the two constituent types can be conjoined by juxtaposition, the language is classified as differentiating. This is actually quite common. In particular, many languages have an overt conjunction marker for nominal conjunction but use juxtaposition for verbal conjunction (e.g. Hausa).

Thus, the map shows the distribution of the following three language types:

1.	Nominal and verbal conjunction are largely identical	161
2.	Nominal and verbal conjunction are different	125
3.	Nominal and verbal conjunction are primarily expressed by juxtaposition	15
	total	301

In illustrating conjunction markers, I will sometimes use the variables *A*, *B*, etc. for nominal conjuncts, and the variables *P*, *Q*, etc. for verbal conjuncts (including clauses). For instance, Dagbani is said to have the markers *A mini B* and *P ka Q*.

2 Multiple marking strategies

It is of course not uncommon for languages to possess several ways of conjoining noun phrases, and/or several ways of conjoining verb phrases and clauses. This is unproblematic when the competing markers have exactly the same range of uses. For instance, Madurese (Sundic, Austronesian; Java, Indonesia) has the two conjunction markers *ban* and *biq*, and both of these are used for nominal and for verbal conjunction (*A ban B, A biq B, P ban Q, P biq Q*; Davies 1999: 25). Another example is Wardaman (Yangmanic; Northern Territory, Australia), where there are two different markers for nominal conjunction (*A warrma B*, and *A B wayana*) and one marker for verbal conjunction (*P gabani Q*) (Merlan 1994: 68, 87, 89). In contrast to Madurese, where there is complete overlap between nominal and verbal conjunction markers, Wardaman shows no overlap at all; hence Madurese is classified as "largely identical", and Wardaman is classified as "different".

But quite a few languages show **partial overlap** of their nominal and verbal conjunction markers. In such cases, the language is generally classified as "largely identical". For example, Southeastern Tepehuan (Southern Uto-Aztecan; Durango, Mexico) has two conjunction markers, *gam* and *guio*. *Gam* only links noun phrases describing entities of a similar nature, while *guio* links both noun phrases and clauses (Willett 1991: 216–17). And in Tinrin (Oceanic; New Caledonia), *mwâ* only conjoins clauses (verb phrases cannot be conjoined in Tinrin), whereas *mê* conjoins both clauses and phrases. The difference between *mwâ* and *mê* in clause conjunction is that *mê* is used when the two clauses are parallel and *mwâ* is used when they contrast in some way (Osumi 1995: 258).

However, cases of partial overlap are classified as "different" under two circumstances: (i) if the overlap concerns a marker which is a clearly minor alternative for at least one of the constituent types, and (ii) if the overlap concerns the juxtaposition strategy. Thus, Koromfe (Gur, Niger-Congo; Burkina Faso and Mali) was classified as showing different conjunction markers, although its nominal marker *la* (*A la B*) can also be used for verbal conjunction (*P la Q*). However, juxtaposition is by far the more common means of verbal conjunction in the language, and when it combines two clauses, *la* mostly means 'but' rather than 'and' (Rennison 1997: 88, 92). Thus, *la* was judged to be a minor alternative for verbal conjunction. Similarly, in Imonda (Border family; Papua New Guinea), the major alternatives for nominal conjunction are *A-i B-i* (for human conjuncts, e.g. *ehe-i ka-i* 'he and I') and *A-na B-na* (e.g. *sa-na fo-na* 'coconut and banana'). For verbal conjunction, the major alternative is juxtaposition, and there are two other markers: *P-ie Q-ie*, and *P-mo Q*. The element *-mo* also occurs in a minor nominal conjunction type, *A-mo B-mo C-mo . . .*, which is used for enumeration (Seiler 1985: 68–70, 102, 196–7). Thus, *-mo* occurs both in nominal and in verbal conjunction, but not as a major option in either, so it was disregarded.

Juxtaposition is extremely common as an alternative option, especially for clausal conjunction, but also for nominal conjunction.

It is therefore not very telling if a language uses it for both constituent types; hence if there are also overt conjunction markers, only these were taken into account. For example, in Trumai (isolate; Mato Grosso, Brazil), juxtaposition is a possible means of conjoining both noun phrases and verb phrases/clauses, but there are two other nominal conjunction markers (*A B a*, *A B tam*) and several other verbal conjunction markers (e.g. *P inis Q*) (Guirardello 1999: 19–21, 367–70), so Trumai was classified as "different".

3 Identifying conjunction

Conjunction is a subtype of **coordination**. We are dealing with a coordinating construction when there are two identifiable constituents which have the same semantic role and together form a larger constituent (Haspelmath 2005). Conjunction is the type of coordination that is translated by 'and', as opposed to **disjunction** ('or') and **adversative coordination** ('but').

There are usually few problems in recognizing nominal conjunction. Difficulties may arise when the conjunction marker is the same as the **comitative** adposition (or case affix) 'with', as is the case in a great many languages (see Chapter 63). Consider, for instance, the case of Babungo (Grassfields Bantoid, Niger-Congo; Cameroon), where *nə̀* is translated as 'and' or 'with'. However, it is quite clear that *nə̀* in (4) is a conjunction marker:

(4) Babungo (Schaub 1985: 84–5)
Lámbí ***nə̀*** *Ndúlà* ***nə̀*** *Kɔ̀mə̀tá gə̀ táa yìwìŋ.*
Lambi and Ndula and Kometa go-PFV to market
'Lambi, Ndula, and Kometa went to the market.'

There are four reasons here why *nə̀* 'and' cannot be simply identical with *nə̀* 'with': It occurs within the subject noun phrase rather than after the verb (as adverbial prepositions do); it occurs twice and links three NPs to each other; it need not imply that the referents carried out the action jointly; and finally, *nə̀* 'with' no longer has the comitative meaning which must have been the original source of the conjunction meaning—it is now confined to the instrumental meaning 'with' (e.g. *nə̀ fə̀ntɨ̄* 'with a stick'; Schaub 1985: 145). Although conjunction markers are often similar to comitative adpositions because they tend to develop from them diachronically, the two are in general fairly easy to tell apart. Whenever a marker is part of a noun phrase and is systematically translated as 'and' or described as a coordinating conjunction, it is regarded as a conjunction marker here.

Conjunction of clauses is more difficult to identify. On the one hand, it is not always clear whether we are dealing with a sequence of separate sentences or a complex sentence consisting of two conjuncts. If there are no overt markers, the best general criterion for recognizing conjunction is intonation, but since this is not described systematically in grammars, it is sometimes necessary to rely on the translation: when a clause-combining strategy is systematically rendered by 'and', then it is regarded as conjunction.

On the other hand, it is often difficult to distinguish between verbal/clausal conjunction and verbal/clausal **subordination**. In many languages, the most natural translation of 'A did X and B did Y' is by means of a special **dependent verb form** (sometimes called **converb** or **medial verb**) in one of the clauses. Examples come from Japanese and Tauya (Madang, Trans-New Guinea; Papua New Guinea).

(5) Japanese (Kuno 1978: 123–4)
*Taroo ga Amerika ni ik-**i**, Hanako ga*
Taro NOM America DAT go-CONV Hanako NOM
Furansu ni it-ta.
France DAT go-PST
'Taro went to America, and Hanako went to France.'

(6) Tauya (MacDonald 1990: 219)
*Yate-e-**te** ni-a-ʔa.*
go-1-MED eat-3SG-IND
'I went and he ate.'

The verb forms marked by *-i* in Japanese and *-te* in Tauya do not occur in independent clauses, so in that sense the first clauses in (5–6) could be regarded as subordinate, and one could claim that a more literal translation of (5) would be 'Taro having gone to America, Hanako went to France'. However, when other criteria for the subordination distinction are taken into account, these sentences do not come out consistently as subordinate (see Haspelmath 1995 for general discussion, and Alpatov and Podlesskaya 1995 for Japanese). They thus show that the subordination/coordination distinction cannot be drawn in a straightforward way.

For the purposes of this chapter, a clause-combining strategy is regarded as a type of conjunctive strategy if it is consistently translated by 'and' and not by more concrete adverbial markers (such as 'after', 'while', or 'and then'). This is again a fairly crude criterion, but many grammars do not provide much more information. For a number of languages, no information about conjunction of verb phrases and clauses was available, but the nominal conjunction markers are described as "confined to noun phrase conjunction", so that it could be inferred that the nominal and verbal conjunction markers must be "different". Note that I assume that all languages have some way of expressing both nominal and verbal conjunction.

4 Verb phrase conjunction versus clause conjunction

Some languages use different markers for verb phrase conjunction and clause conjunction. For instance, Somali has three different markers: *A iyo B* for nominal conjunction, *P oo Q* for verb phrase conjunction, and *P Q-na* for clause conjunction:

(7) Somali (Berchem 1991: 324–7)
a. *rooti* ***iyo*** *khudrat*
bread and fruit
'bread and fruit'
b. *Wuu cunay* ***oo*** *cabbay.*
FOC.3M.SG eat and drink
'He ate and drank.'
c. *Macallin-ku wuxuu joogaa dugsi-ga,*
teacher-ART FOC.3M.SG be school-ART
*carruur-ta-**na** waxay ku cayaarayaan*
children-ART-and FOC.3PL PREV play
dibed-da.
outside-ART
'The teacher is in the school, and the children are playing outside.'

Somali is straightforwardly classified as showing "different" nominal and verbal conjunction. But in a few languages, the marker for verb phrase conjunction is identical to the nominal conjunction marker, but different from the clausal marker. Such a language is Chamorro (Western Malayo-Polynesian; Guam).

(8) Chamorro (Topping 1973: 146)
a. *si Juan* ***yan*** *si Maria*
ART Juan and ART Maria
'Juan and Maria'
b. *Kumakati* ***yan*** *chumachalek i palao'an.*
cry and laugh ART woman
'The woman is crying and laughing.'
c. *Malago' yo' ni lepblo* ***ya*** *hu fahan.*
want I ART book and I buy
'I wanted the book and I bought it.'

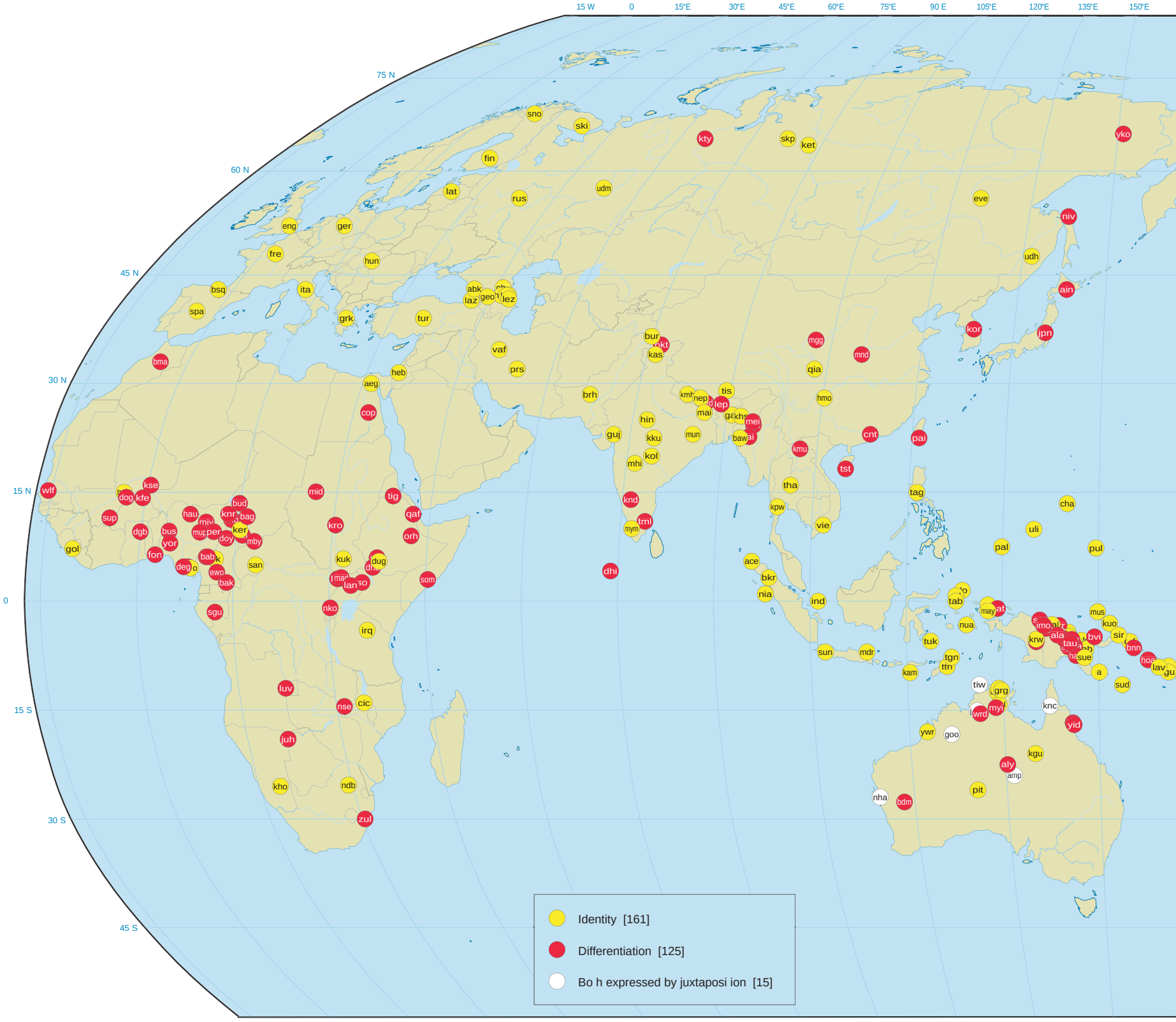
In such cases (which are quite rare), nominal conjunction is compared with clausal conjunction, so that these languages are classified as "different".

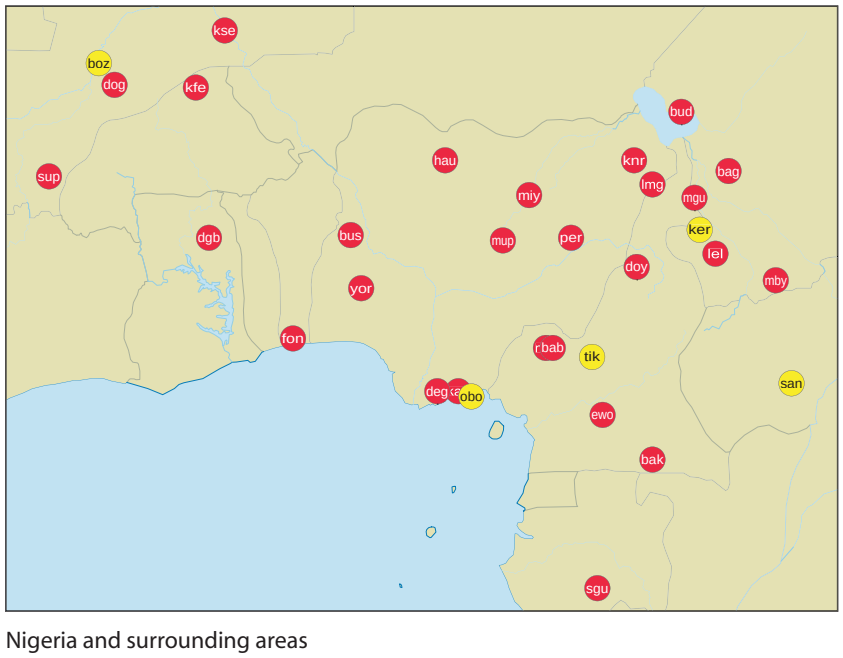
5 Geographical distribution

Differentiation of nominal and verbal conjunction markers is particularly widespread in African languages (as was noted by Welmers 1973: 305), but it is also found in many other parts of the world. Identity of the two markers is particularly common in Europe and Mesoamerica, but it is also found widely in the rest of the world. Juxtaposition is particularly common in Australia and South America.

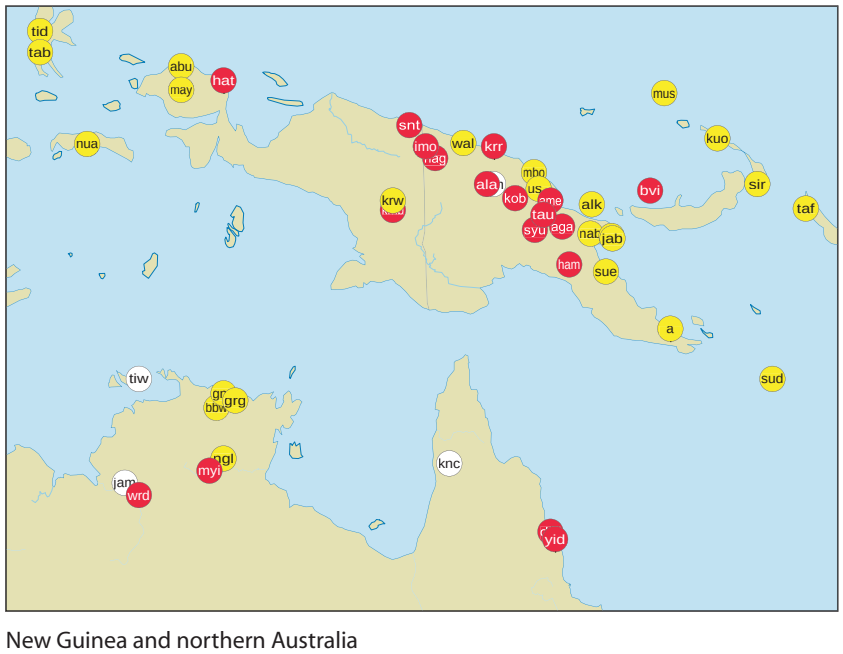
64 Nominal and Verbal Conjunction

MARTIN HASPELMATH

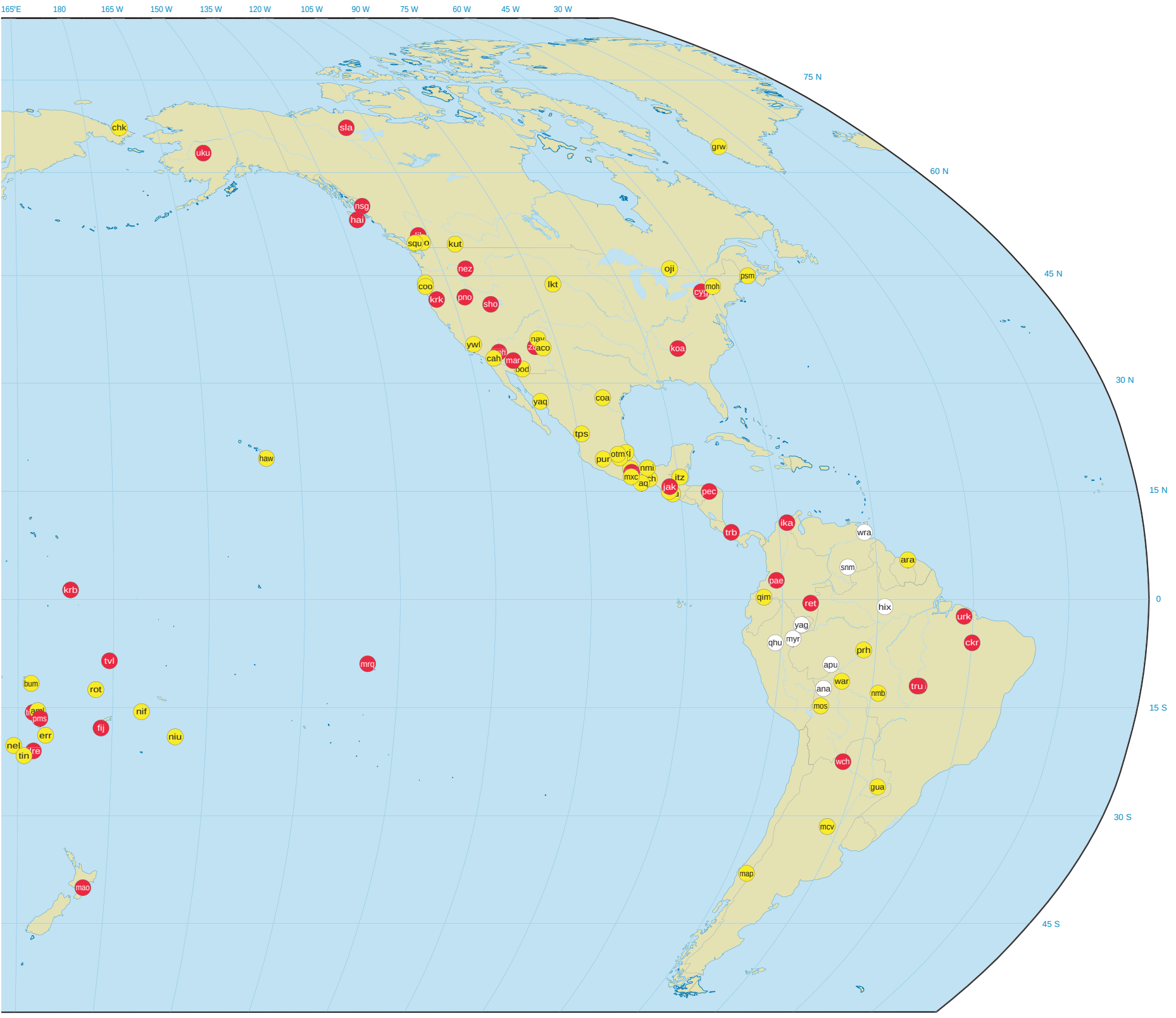


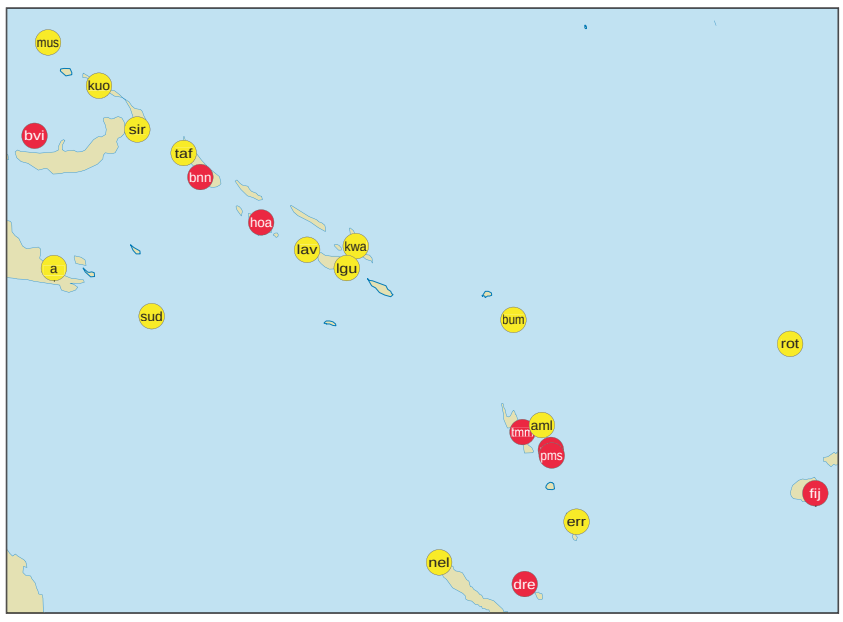


Nigeria and surrounding areas

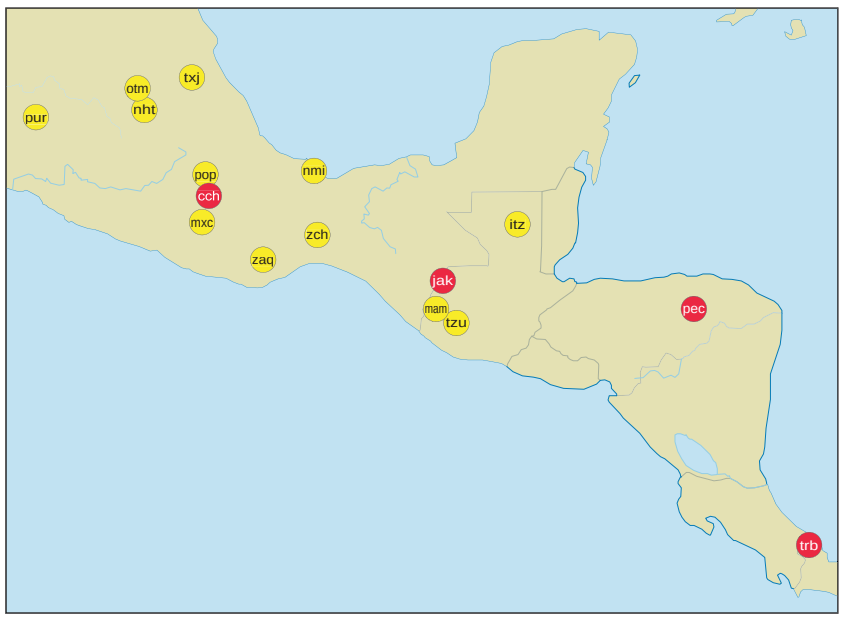


New Guinea and northern Australia





islands east of New Guinea



southern Mexico and Central America

65–68 Tense and Aspect

ÖSTEN DAHL AND VIVEKA VELUPILLAI

General introduction

The following four maps concern the role of **tense** and **aspect** in the grammars of the world's languages. Traditionally, tense and aspect are seen as **grammatical categories** of verbs. Indeed, insofar as verbs display morphological variation in languages, there is a strong tendency for such variation to include inflectional differences between forms that reflect temporal and aspectual distinctions, such as the difference between the present tense form *sings* and the past tense form *sang* in English. However, some modifications to the identification of tense and aspect with grammatical categories of verbs must be made. To begin with, it has long been recognized that **periphrastic constructions**, such as the English Perfect (*I have sung*) and Progressive (*I am singing*), are employed in functions similar to those of inflections. Furthermore, tense and aspect do not always present themselves as separate and neatly delineated categories, as we shall now argue.

Among the multitudinous definitions of tense and aspect in the literature, we may cite those given in Comrie (1985: 1, 6): tense is "grammaticalisation of location in time" and aspect is "grammaticalisation of expression of internal temporal constituency" (of events, processes etc.). Thus defined, the two categories are conceptually close in that both deal with time. But they may also be interwoven in grammatical systems in that one and the same grammatical form may combine temporal and aspectual elements in its semantics. For instance, the very common and central distinction between imperfective and perfective verb forms would appear to be a straightforward example of aspect; yet it typically involves not only aspectual notions but also restrictions on temporal reference, in that perfective verb forms are usually taken to refer to past events. This fact has a cognitive basis: we may say that the prototypical uses of perfectives coincide with the default view of an event as a completed whole. But normally such a perspective is possible only if the event is wholly in the past. Therefore, in languages which do not mark the imperfective-perfective distinction, verbs that are typically event-denoting such as 'arrive', 'close', and 'kill' may be interpreted as referring to the past in the absence of any marking to the contrary. Verbs that typically denote states such as 'sleep', on the other hand, tend to be interpreted as referring to the present. (In addition, verbs are sometimes indeterminate or ambiguous between referring to a present state or to the event in the past that gave rise to it, e.g. German *er ist gestorben* 'he is dead' or 'he died'.)

What has been said here about tense and aspect may also be extended to the closely related category of **mood**. An alternative to seeing tense, aspect, and mood as grammatical categories in the traditional sense is to regard tense-aspect-mood systems as wholes where the building blocks are the individual tenses, aspects, and moods, such as the Past and the Progressive in English. These will be referred to as **grams**, and it is assumed that on the cross-linguistic level they represent a restricted set of **gram types**. This is the approach that will be taken here. In the following, the names of language-specific grams will be capitalized, to distinguish them from gram types, which are written in lower-case characters.

The term **grammaticalization**, as used in the definitions from Comrie cited above, refers to a synchronic property characterizing a notion (semantic category) if and only if it is reflected in or determines the use of grammatical items. Obviously, both location in time and the internal constituency of events may receive linguistic expression in multifarious ways, which are not, however, relevant to the notions of tense and aspect as defined here, as long as they do not play a systematic role in the grammar. For instance, temporal adverbials such as *yesterday* or *in 2001* are also used to locate events in time, but they differ from grammatical tense in fundamental ways. While temporal adverbials are used only when they express information which is relevant to the particular intended message, the use of tenses is guided by general principles that often make the choice of a certain tense **obligatory** and that make the use of a tense morpheme obligatory even if the information it carries is redundant. Thus, in a sentence such as *Last year I bought a new car* the choice of a tense other than the simple past would make the sentence anomalous, although the information that the event took place in the past is expressed unambiguously by *last year*.

In the case of aspect, the delimitation of what is grammatical(ized) is more difficult, and further confounded by the lack of a consistent terminology. In some research traditions, the term "aspect" is used to refer to a wide domain of phenomena, including many that are not manifested as grammatical distinctions (cf. for instance Verkuyl (1971), Tenny (1994)). Others make a strict distinction between aspect as a grammatical phenomenon and "Aktionsart" as pertaining to lexical or purely notional (semantic) categories (this tradition goes back to Agrell (1908)). However, among those who make a distinction between aspect and Aktionsart, there is no unanimity as to how the latter term should be used. Especially in Slavic linguistics, the term "Aktionsart", or its counterpart in other languages (such as *sposob dejstvija* in Russian), is used for phenomena that straddle the borderline between grammar and lexicon, notably various derivational processes by which verbs with specific aspectual meanings may be created. For instance, from the simplex Russian verb *spat'* 'to sleep' we may obtain the verb *pospat'* 'to sleep for a while' by adding the "delimitative" prefix *po-*. With a more liberal definition of aspect, such processes would be called "derivational aspect". Although there is considerable variation as to the extent to which derivational aspect is elaborated and used in languages, relatively little work has been done with the aim of systematizing this variation. In these chapters, we have therefore chosen to focus on grammatical aspect in the narrowest sense, where typological research has already advanced far enough to make it possible to define mappable parameters. However, some areal tendencies may be mentioned here. Derivational aspect (Aktionsart) appears to be particularly well developed in many indigenous North American languages; eastern Europe could be another example of an area where these phenomena are well represented.

Today, "grammaticalization" is probably most often understood in a diachronic sense—as the development of grammatical marking over time, typically from lexical sources. The grammaticalization paths for tense and aspect are relatively well known, at least in the languages for which there is sufficient information about earlier stages (which is certainly not a representative sample of the world's languages). It has thus, for instance, been established that both pasts and perfectives may arise from perfects, whereas imperfectives often develop out of progressives (see Bybee et al. 1994: 51–175); iteratives and similar constructions as well as futures may come from a number of different sources, such as verbs of volition, obligation, and motion (Bybee et al. 1994: 243–80). There is a strong correlation between the way in which a tense-aspect gram is expressed —whether it is inflectional or periphrastic—and how far it has advanced on its path of grammaticalization. Thus, perfects and progressives are overwhelmingly periphrastic, whereas pasts and perfectives are more prone to be inflectional.

It goes without saying that it is impossible to encompass the richness of the tense and aspect systems of the world's languages within four maps. We have been forced to neglect a number of gram types, most of which have interesting geographical distributions; among these are habituals, iteratives, frequentatives, dedicated narrative forms, resultatives, and experientials. We have also neglected tense and aspect marking in embedded contexts, as well as the interplay

between tense-aspect and other categories, such as negation (many languages have special tense-aspect forms in negative contexts, cf. Chapter 114).

Tense and aspect are notoriously difficult categories to describe adequately, and the treatment in grammars is often problematic, especially if one wants to use it for cross-linguistic comparison. As far as possible, we have tried to apply consistent criteria in classifying tense-aspect phenomena. For this reason, our interpretations sometimes differ from those found in grammars. The reader should thus not be surprised if a language is classified in an unexpected way.

In compiling the data for the maps on tense and aspect, we have drawn upon two earlier large-scale typological surveys, namely those presented in Dahl (1985) and Bybee et al. (1994). The sample is the same for all four maps and contains all languages in the basic *WALS* 100-language sample except one (Mezquital Otomí), for which it was not possible to get adequate information. Discussions of areal tendencies in tense-aspect systems based on the sample in Bybee et al. (1994) are found in Dahl (1995) and Dahl (2000); the present sample is about three times as large, however, and although earlier findings are basically confirmed, we are now able to see the general pattern much more clearly.

65 Perfective/Imperfective Aspect

1 Introduction

The distinction between **imperfective** and **perfective** plays an important role in many verb systems and is commonly signalled by morphological means (rather than being expressed periphrastically). A particularly straightforward case is found in Rendille (East Cushitic; Kenya). Nonstative verbs in Rendille distinguish two basic forms, one which normally ends in *-a* and one which normally ends in *-e*, as illustrated by the examples in (1).

(1) Rendille (own data)

a. *khadaabbe* *chiirta*
letter.PL write.IMPF
'He writes/is writing/wrote/was writing/will write letters.'

b. *khadaabbe* *chiirte*
letter.PL write.PFV
'He wrote letters.'

The imperfective form in *-a* is used for reference to the present and the future but also for ongoing and habitual events in the past, as indicated by the translations. The perfective form in *-e* is basically restricted to single completed events in the past (with some vacillation for past habitual contexts). In most other languages with an imperfective/perfective distinction, this pattern is obscured by interaction with other tense-aspect grams, but the basic opposition between one form (or set of forms) which is used exclusively or almost exclusively for single completed events in the past and another form (or set of forms) which is used for everything else is characteristic of the distinction.

To be interpreted as a perfective, we demand that a form should be the default way of referring to a completed event in the language in question. In many languages, there are forms or constructions that are used of completed events but only if some additional nuance of meaning is intended, for instance if emphasis is put on the result being complete or affecting the object totally. Such **strong perfectives** ("conclusives" in Dahl (1985) and "completives" in Bybee et al. (1994)) exhibit relatively large variation cross-linguistically. They are often called "perfectives" in grammars but are not counted as such here. The following example is from Rama (Chibchan; Nicaragua), where the suffix *-atkul-* (derived from a verb meaning 'finish') indicates a "strong perfective":

(2) Rama (Grinevald n.d.: 154)
yaing *kwiik* *alauk-atkul-u*
3.POSS hand burn-STRONG.PFV-PST
'He burned his hand completely.'

We distinguish imperfectives from **progressives**, with which they partially overlap and which are often seen as a variety of imperfectives. Progressives, like the English *is singing* or the equivalent Spanish *está cantando*, have a more restricted domain of use (for instance, they are typically not the primary choice for expressing habitual meaning), which means that they are opposed to nonprogressive forms independently of time reference. They are also normally restricted to nonstative verbs. Progressives are frequent diachronic sources of marked imperfectives, and borderline cases admittedly exist.

We also distinguish imperfectives from **antipassives**, by which we understand processes that operate on transitive constructions to make them less transitive by removing the direct object or marking it as an oblique (see Chapter 108). Antipassives, which are particularly frequent in ergative languages, often influence the aspectual character of the sentence, with ranges of meaning similar to those of progressives and imperfectives. The following example is from Bandjalang (Pama-Nyungan; New South Wales, Australia):

(3) Bandjalang (own questionnaire data from M. J. Sharpe)
nyule *leda* *bugalehn*
he.VISIBLE letter write.ANTIPASSIVE
'He is writing a letter.'

2 Definition of values

For this map, only two values have been defined: languages in which there is grammatical marking of the perfective/imperfective distinction (red) and those where there is not (white). "Grammatical marking" here includes both marking by morphological means and by periphrastic constructions.

In most other tense-aspect oppositions, there is cross-linguistic consistency with respect to which member of the opposition is to count as the marked one. For the imperfective/perfective distinction, this is not possible. There are languages in which the perfective has no marker and the imperfective has an overt marker, and vice versa, but most often (at least in our sample) no clear marking relations can be identified. (One reason for this is that the distinction is frequently manifested by stem alternations and similar processes.)

1.	Grammatical marking of perfective/imperfective distinction	101
2.	No grammatical marking of perfective/imperfective distinction	121
	total	222

3 Geographical distribution

Perfective/imperfective distinctions seem to be less skewed in their geographical distribution than, for instance, past tenses (see Chapter 66). However, we can discern the following tendencies. In a band across southern Eurasia from Europe (excluding most of the northern part) to China (but excluding the Dravidian part of South Asia and all of South-East Asia), there is fairly consistent marking of perfectivity/imperfectivity. One may see this area as extending into Africa down to the Equator. Interestingly, it overlaps quite

considerably with the Eurasian/African past-marking area (seen on Map 66), but lies south of it. Northern Europe outside the Slavic area has very little perfectivity/imperfectivity marking. Other white clusters on the map include large parts of South America and South-East Asia (see remarks on the latter in Chapter 66).

4 Further considerations

Even if perhaps not so often formulated as an explicit hypothesis, there seems to be a widespread view of tense and aspect as alternatives to each other—that languages tend to be either "tense languages" or "aspect languages". If this were the case, we would expect a negative correlation between imperfectives and perfectives on the one hand, and pasts and futures on the other. The data presented here provide no support for such a conclusion. In fact, there are considerably more languages in the sample that have both the aspectual and the temporal categories, or neither of the alternatives, than have one only. It is plausible that there is rather a positive correlation between all the categories under discussion and the general morphological complexity of the verb.

66 The Past Tense

1 Introduction

In English, like virtually all European languages, there is a systematic grammatical distinction between **present tenses** and **past tenses**, as in the following sentence pair:

(4) a. *The temperature is below zero right now.*
b. *The temperature was below zero yesterday at noon.*

In (4) the form of the finite verb (in this case, the copula *is*/*was*) depends on what time we are talking about—what we may call the **topic time** (following Klein (1994)): the present form *is* is used if the topic time coincides with the time of speech, and the past form *was* if the topic time precedes the time of speech. This is of course a very rough rule in need of a number of further specifications. As was already noted in the introductory chapter, grammaticalized marking of time in the form of tenses is typically independent of considerations of relevance. The fact that time in (4) is also indicated (and more precisely so) by time adverbials does not make tense marking less necessary. Rather, the presence of a deictic time adverbial such as *yesterday* renders the use of anything but the simple past tense unacceptable in (4b).

Since it is generally the past tense rather than the present that is overtly marked, we may speak of languages having or not having past marking rather than having a past/non-past distinction.

It is only from a Eurocentric point of view that the marking of the distinction between present and past appears to be a necessary part of grammar. Languages may or may not distinguish (4) grammatically, and there is no clear majority for either alternative. The following Indonesian example, which translates both 'The water is cold' and 'The water was cold', illustrates the lack of a present/past distinction:

(5) Indonesian (own data)
Air itu dingin.
water that cold
'The water is/was cold.'

The primary goal of this chapter is to show the geographical distribution of past tenses. There are several things that make this task less straightforward than it might seem prima facie.

Perhaps most importantly, we must elucidate the relationship between past tenses and grammatical aspect. Given that perfective forms are by default interpreted as referring to the past (see the general introduction to Chapters 65–68), any further marking of past time reference would appear redundant. Indeed, it is probably most common for overt past-tense marking not to be compatible with perfective aspect. In such cases, we may get the **tripartite** system represented, for example, by the present-imperfect-aorist system of many Indo-European languages, illustrated here by Eastern Armenian, where a periphrastic construction (copula + converb of the main verb) is used in the imperfective (Present and Imperfect tenses below):

(6) Eastern Armenian (own data)
a. Present:
Na namak e gər-um.
he letter is write-CONV
'He is writing/writes a letter.'
b. Imperfect:
Na namak er gər-um.
he letter was write-CONV
'He was writing/wrote (habitually) a letter.'
c. Aorist:
Na namak gr-ecə.
he letter write-AOR.3SG
'He wrote a letter (a single event).'

Similar systems are found in many different families, although they are particularly common in the past-marking part of Eurasia. The question is how they should be analysed. The possibility that first comes to mind is probably to assign the labels "present", "past imperfective", and "past perfective" to the three forms. This would imply that (6b) and (6c) are both members of a general past-tense category, and that the aspectual distinction is made only there. However, there are at least two arguments against such an analysis. One is that it does not fit the actual make-up of the forms very well. In the Eastern Armenian system, there is nothing that unites (6b) and (6c) in the way they are expressed. Rather, (6b) is naturally seen as an elaboration on (6a), and (6c) has a wholly unique composition. This state of affairs turns out to be typical of tripartite systems (although some exceptions can be found). Another argument is that calling (6c) a "past perfective" obscures the basic cross-linguistic unity of perfectives, which appear to have more or less the same semantics irrespective of whether the language distinguishes past and present in the imperfective or not. Following Dahl (1985: 81–4), we therefore do not regard categories like the Armenian Aorist as "past perfectives" but rather simply as "perfectives", and categories like the Armenian Imperfect as a particular variety of past, restricted to imperfective forms. In Dahl (1985: 117–18), such forms were labelled "PASTi" but here we use the more intuitive "past imperfective". Aspectually unrestricted pasts and past imperfectives are not distinguished on the map, the argument being that the latter are found in the overwhelming majority of languages which have both the perfective/imperfective distinction and past-tense marking. The most notable exceptions are languages such as Russian, where these two grammatical phenomena are more independent of each other.

Both Dahl (1985: 117) and Bybee et al. (1994: 82) provide support for the claim that significantly more than half of all pasts are marked morphologically. The tendency may be weaker for past imperfectives, but the material is not extensive enough to make any certain claims.

Very often, languages make further grammatical distinctions within the domain of past time reference. Thus, many languages

have **perfects** as separate categories—this will be discussed in Chapter 68. In addition, as many as one fifth of the languages in our sample make **remoteness distinctions**—that is, tense choice is dependent on the temporal distance between the time of speech and the topic time. Remoteness may be more subjectively or more objectively determined; in the latter case, a combination of a "remote" time adverbial with a "non-remote" tense will result in ungrammaticality. (It should be noted that like tenses in general, tenses which distinguish degrees of remoteness do not substitute for adverbials but are used whether or not there is another temporal indication in the sentence.)

Almost universally, if there is one well-defined cut-off point in the past between different forms, the division lies between 'today' and 'before today'. The 'before today' range is often divided further. The term **hodiernal** is commonly used for 'today's past', and tenses that are restricted to the day before the point of speech may be called **hesternal**.

The richest system in our sample is that of Yagua, which according to the available description (Payne and Payne 1990: 386–8) has five degrees of remoteness in the past, as shown in Table 1:

Table 1 Remoteness distinctions in Yagua

Name in grammar	Use	Suffix	Example
Proximate 1	'a few hours previous to the time of utterance'	*-jásiy*	*rayą́ą́siy* *{ray-jiya-jásiy}* 1SG-go-PROX1 'I went (this morning).'
Proximate 2	'one day previous to the time of utterance'	*-jay*	*rį̨ų́núújeñíí* *{ray-jų̃nnúúy-jay-níí}* 1SG-see-PROX2–3SG 'I saw him (yesterday).'
Past 1	'roughly one week ago to one month ago'	*-siy*	*sadíichimyaa* *{sa-díí-siy-maa}* 3SG-die-PST2-PERF 'He has died (between a week and a month ago).'
Past 2	'roughly one to two months ago up to one or two years ago'	*-tíy*	*sadíitímyaa* *{sa-dííy-tíy-maa}* 3SG-die-PST2-PERF 'He has died (between 1 to 2 months and a year ago).'
Past 3	'distant or legendary past'	*-jada*	*raryúpeeda* *{ray-rupay-jada}* 1SG-be.born-PST3 'I was born (a number of years ago).'

Forms used for recent past sometimes coincide with perfects (or are historically derived from them). As the label 'distant or legendary past' in the table suggests, it may sometimes be hard to distinguish distant pasts from forms reserved for use in myths and legends, which may function not only as tenses, but also as stylistic or modal markers.

2 Definition of values

The basic dichotomy here is between languages that mark the past/non-past distinction grammatically (including marking by periphrastic constructions) and those which do not. Within the first group, three subtypes are distinguished, depending on the number of remoteness distinctions made, as shown in the feature-value box.

1.	Past/non-past distinction marked; no remoteness distinction	94
2.	Past/non-past distinction marked; 2–3 degrees of remoteness distinguished	38
3.	Past/non-past distinction marked; at least 4 degrees of remoteness distinguished	2
4.	No grammatical marking of past/non-past distinction	88
	total	222

3 Geographical distribution

There are quite strong areal tendencies in the distribution of past tenses—perhaps strong enough to deserve another word than "tendency". These can be seen on the map as homogeneous one-coloured areas.

The largest homogeneous past-marking area is one that stretches from Iceland in the north-west to the Horn of Africa in the south and Bangladesh in the south-east, including much of central Eurasia (but excluding northeastern Siberia). Since the Indo-European phylum also extends between Iceland and Bangladesh, it is tempting to see this as an Indo-European phenomenon. However, the area also includes phyla such as Uralic, Altaic, and large parts of Afro-Asiatic, and isolates such as Basque and Burushaski. Other homogeneous past-marking areas include Australia, northern South America, and central New Guinea.

In Africa, the majority of the sample languages are past-marking, but there is also an (almost) homogeneous nonmarking area in central and western Africa. Two things are worth mentioning here. One is that most of the languages involved have an imperfective/perfective distinction. The other is that the languages without past in this area include members of all three major phyla that are represented here—Afro-Asiatic, Niger-Congo, and Nilo-Saharan. None of these phyla is consistently nonmarking, however. In other words, we are dealing with a fairly clear example of areal convergence.

The most salient area of homogeneous nonmarking of past is found in South-East Asia, and also includes languages from a number of different phyla—Sino-Tibetan, Austro-Asiatic, and Austronesian, to mention the largest ones. The languages in this area lack not only past tenses but also marking of the imperfective/perfective distinction and inflectional futures. These are of course the well-known isolating languages of South-East Asia, and it is hardly a coincidence that they lack precisely those tense-aspect gram types that cross-linguistically are most often marked inflectionally. What is more difficult to decide is what is cause and what is effect here.

There may be a similar area in West Africa, but it is at any rate much more restricted—in our sample there are only three West African languages that are marked as white on all three of the Maps 65–67.

67 The Future Tense

1 Introduction

In English, the sentence *It is cold tomorrow*, with the present tense of the copula *is*, sounds strange: it is more natural to say *It will (it'll) be cold tomorrow* or *It is going to be cold tomorrow*, using a **future tense** form. In Finnish, on the other hand, one may replace the adverb *tänään* 'today' in (7) with *huomenna* 'tomorrow', yielding (8) without any further changes in the sentence.

Finnish

(7) *Tänään on kylmää.*
today is cold.PART
'It is cold today.'

(8) *Huomenna on kylmää.*
tomorrow is cold.PART
'It will be cold tomorrow.'

The Present Tense can thus be used equally well for the present and the future in Finnish, in contrast to English, where it is often the case that auxiliary constructions such as *shall/will*+Verb and *be going to*+Verb must be used when speaking of the future. In many languages where future time reference is grammaticalized, the means employed is inflectional. Thus, in French, (9) and (10) differ in the form of the verb *faire* '(lit.) do'.

French

(9) *Il fait froid aujourd'hui.*
it do.PRES.3SG cold today
'It is cold today.'

(10) *Il fera froid demain.*
it do.FUT.3SG cold tomorrow
'It will be cold tomorrow.'

It is relatively rare for a language to totally lack any grammatical means for marking the future. Most languages have at least weakly grammaticalized devices for doing so. In this chapter, we have therefore decided to map only the **inflectionally marked future tenses**, inflectional marking being a relatively clear criterion (although there are some borderline cases where it is unclear if one is dealing with a clitic or an affix). Inflectional markings more often tend to be obligatory and also on the whole have a wider range of uses. For instance, they regularly show up in temporal and subordinate clauses, where periphrastic future-marking devices are relatively rare. They also appear systematically (often obligatorily) in sentences which express clear predictions about the future (which are independent of human intentions and planning), whereas less grammaticalized constructions often tend to be predominantly used in talk of plans and intentions—a fact which is explainable from the diachronic sources of future tenses, which have been fairly well studied (Bybee et al. 1994: 243–80). In most cases, inflectional future tenses derive from periphrastic constructions (employing auxiliaries or particles), which are in turn derived from constructions expressing such notions as obligation ('must'), volition/intention ('want'), and motion ('go' and 'come'). However, a future tense may develop out of an earlier non-past or imperfective as an indirect effect, for example of the functional expansion of an earlier progressive—the future uses are what is left of the old category after that expansion.

The modal overtones that tend to go with futures have led many linguists to question their status as tenses (e.g. Lyons 1968: 306–11). In the approach presented in our general introduction, it is generally not expected that one will be able to make an unequivocal classification of the elements of tense-aspect-mood systems into neat compartments. From a diachronic point of view, it may be noted that one result of the progressive grammaticalization of futures is that the temporal component of their semantics becomes more dominant relative to the modal component.

Many grammars subsume grammatical future-marking devices under the heading "irreal(is)", especially when their range of use includes negated sentences, counterfactual conditionals, imperatives, etc. With Bybee et al. (1994: 240), we take the view that the distribution of irrealis categories varies too much across languages for them to be acknowledged as a viable cross-linguistic type; such categories are here counted as inflectional futures, if they are expressed inflectionally and cover the same range of uses as other future tenses.

2 Definition of values

For this map, only two values have been defined: languages in which there is inflectional marking of future time reference and those where there is not.

1.	Inflectional marking of future/nonfuture distinction	110
2.	No inflectional marking of future/nonfuture distinction	112
	total	222

3 Geographical distribution

The map shows some fairly clear areal tendencies in the distribution of inflectional future tenses, although there are no real large-scale homogeneous areas as in the case of past tenses (cf. Chapter 66). On two continents, North America and Australia, languages with inflectional futures are in a clear majority, as also in New Guinea (at least the central parts). The sample languages from the South Asian subcontinent consistently mark future inflectionally; this is in stark contrast to the adjacent South-East Asian area, where no inflectional futures are found between 90°E and 120°E. (For a discussion of the South-East Asian area, see Chapter 66.) European languages are averse to inflectional future marking, with some exceptions (the western Romance, Baltic, and Celtic languages). There is a fairly homogeneous future-marking area extending from the Middle East up into the Caucasus. South America and Africa are more varied, although there may be local patterns not visible in this sample.

4 Further considerations

In Chapter 65, we noted that there is no evidence in the data for a division of languages into tense-prominent and aspect-prominent. Another such proposed typology (Ultan 1978) is into languages that oppose grammatically past and non-past and languages that oppose future and nonfuture. Again, the data presented here do not lend support to this typology. There are more languages in the sample that have both pasts and futures, or neither of them, than languages that have only one of the two categories. Even if the proportion may not be wholly reliable, it is unlikely that a negative correlation between the marking of past and the marking of future could be found.

Acknowledgement

We are grateful to Joan Bybee for putting unpublished data from the GRAMCATS database at our disposal.

68 The Perfect

1 Introduction

By **perfect** we mean a category with approximately the same semantics as the English (Present) Perfect in *I have read this book*, which is used to express events that took place before the temporal reference point but which have an effect on or are in some way still relevant at that point. This includes at least two related but distinguishable uses:

(i) A perfect may be used **resultatively**, i.e. of an event, often but not always a recent one, which has results that hold at the time of speech (or any other time serving as reference point): *Someone has stolen my purse!* (=the purse is gone).
(ii) A perfect may be used **experientially**, i.e. to say that a certain type of event took place one or more times over an interval of time, typically one that extends up to the moment of speech (or whatever time serves as the reference point): *I have seen worse things in my life*.

Constructions that have only one of these two uses (dedicated resultatives and experientials) have not been treated as perfects here. Admittedly, the information given in grammars is not always specific enough to make reliable decisions on this point.

Perfects may have further uses, such as the **universal perfect** (or **perfect of persistent situation**), as in *I have lived here for five years*. It is also common for perfects to develop **evidential** uses (see Chapter 77; in English this is typical of the perfect progressive —*You have been drinking*) and **recent past** uses (as in several Romance languages).

Note that the terms "perfect" and "perfective" are not synonymous (the terminology is confusing but has a historical motivation). Bybee et al. (1994) use the term "anterior" in the same sense as "perfect" is used here.

The English Perfect is cross-linguistically typical in being expressed periphrastically.

2 Definition of values

The basic distinction is between languages that have perfects and those that do not. As noted above, only constructions or forms that have both resultative and experiential readings are regarded as perfects here. On the other hand, to count as a perfect, a construction or form must not be equivalent to a general past tense. An operational criterion for judging about this is whether or not the form or construction is regularly used in narratives (Lindstedt 2000: 366).

The most interesting areal patterns with regard to perfects are related to their diachronic sources. Perfects derive diachronically from at least three types of sources, the last two of which are singled out for special marking on the map:

(i) dedicated resultative constructions, usually consisting of constructions involving a past participle (or similar form) in predicative position with or without a copula (sometimes called *esse* perfects or 'be'-perfects), e.g.

(11) Finnish
Juna on saapunut.
train is arrive.SUPINE
'The train has arrived.'

(ii) possessive constructions such as *I have two letters written* (sometimes called HABEO or 'have'-perfects), e.g. the English Perfect;
(iii) constructions involving words such as 'already' or 'finish', as exemplified by Yoruba:

(12) Yoruba (own data)
Ó ti ka iwe na.
he PFV/already read book this
'He has read this book.'

It may be, as argued by Ebert (2001), that what we here call perfects of type (iii) should be treated as a gram type of their own (Ebert calls it "NEWSIT"), where the primary function is to introduce a "new situation". There is also some evidence that a NEWSIT and a perfect, with distinct forms, may coexist in the same language (Burmese and Fijian are possible examples in our sample). We have not tried here to make a systematic distinction between perfects and NEWSITs, since the information at hand does not allow this, but languages with perfects that are known to be diachronically derived from (or have the same shape as) 'already' or 'finish' are marked as such on the map.

There may be further distinctions between perfects in the narrow sense (also called present perfects) on the one hand, and **past perfects** (or **pluperfects**) and **future perfects** on the other. For reasons of space, these are not represented on our maps. It may be noted that, in general, past perfects indeed tend to be realized as combinations of pasts and perfects, and that they are to be expected in those languages in which both these gram types appear. However, sometimes a past perfect may survive the demise of a present perfect, as has happened for instance in Romanian. Past perfects have a relatively strong tendency to develop noncompositional readings; that is, they become semantically independent of pasts and perfects. Future perfects tend to play a more peripheral role in tense-aspect systems.

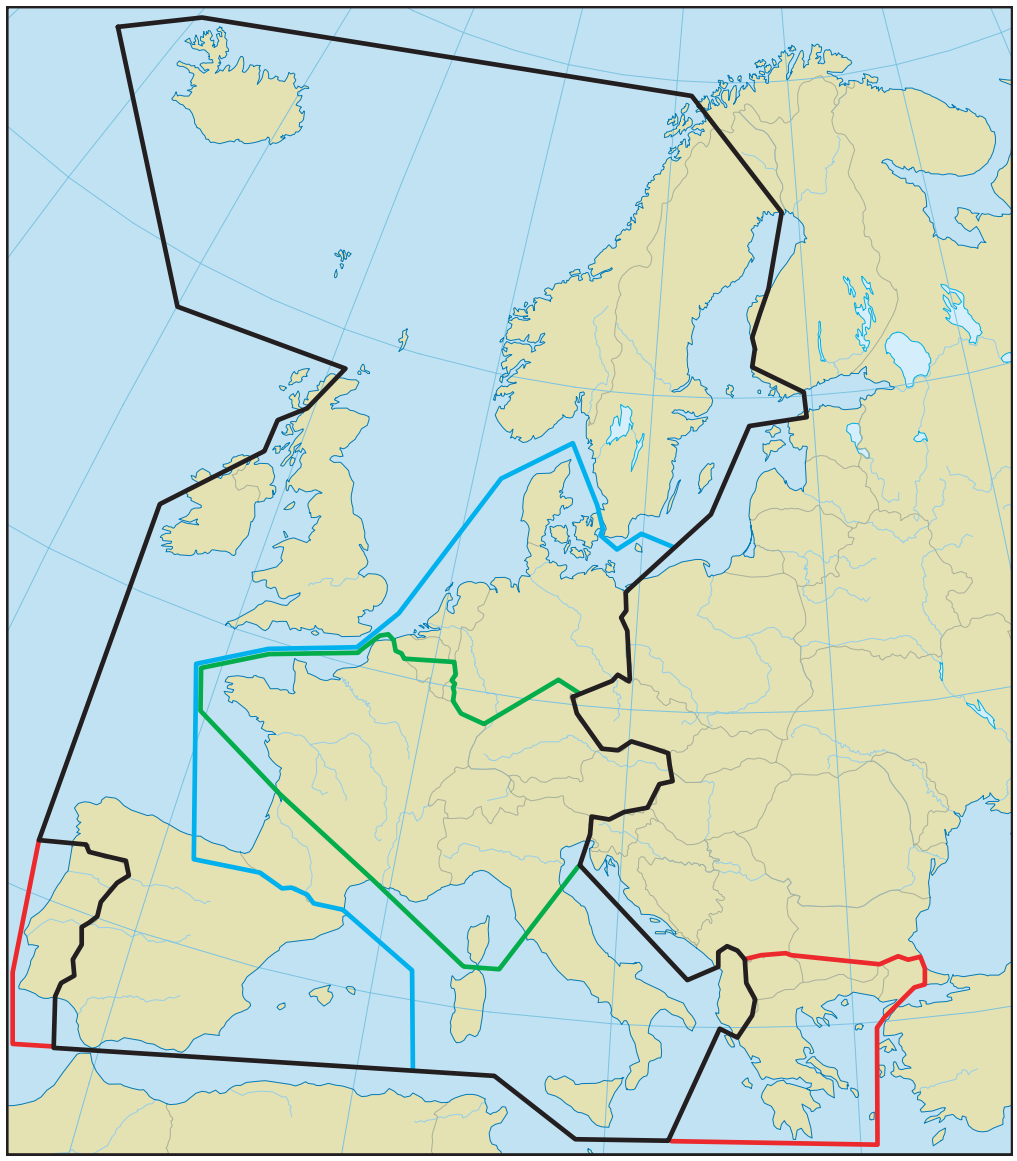
We distinguish three types on the map: perfects known to derive from a possessive construction ('have'-perfects), perfects known to derive from words meaning 'finish' or 'already', and all other perfects (including both perfects known to derive from dedicated resultative constructions, and perfects where the diachronic source cannot be determined).

1. Perfect of the 'have'-type (derived from a possessive construction)	7
2. Perfect derived from word meaning 'finish' or 'already'	21
3. Other perfect	80
4. No perfect	114
total	222

3 Geographical distribution

If we look at the general distribution of perfects, we can note the following tendencies. The largest homogeneous areas with perfects are found in western Europe and South and South-East Asia. A high proportion of languages with perfects is also found throughout Africa and in an area comprising Mesoamerica and the northwestern corner of South America. Large areas with virtually no perfects, on the other hand, are found in the rest of South America and Australia.

However, the differences between the different parts of the world are accentuated when we take into consideration the possible diachronic sources of perfects. Perfects deriving from possessive constructions are attested almost exclusively in Europe (Chukchi is a possible exception, but we have not treated the construction in question as a perfect). Furthermore, they are almost exclusively based on constructions with a transitive possessive verb like English *have*, and are restricted to a contiguous area in western Europe, as a result of an apparently rapid spread in the Middle Ages. Map 68.1 shows the maximal extent of this spread. In the centre of the area, there is a division of labour between 'have'-perfects and 'be'-perfects, the latter being used primarily for intransitive verbs of motion and change. In the centre, there has also been a new development in that the original perfects have come to be used as general pasts or perfectives. The boundaries of these areas are also shown in Map 68.1.



Map 68.1 ‘Have’ perfects in Europe

Present-day boundaries of the area where ‘have’ perfects (perfects based on a transitive possessive construction) and constructions derived from them occur

Peripheral area where less typical ‘have’ perfects are found

Area of division of labour between ‘have’ perfects and ‘be’ perfects

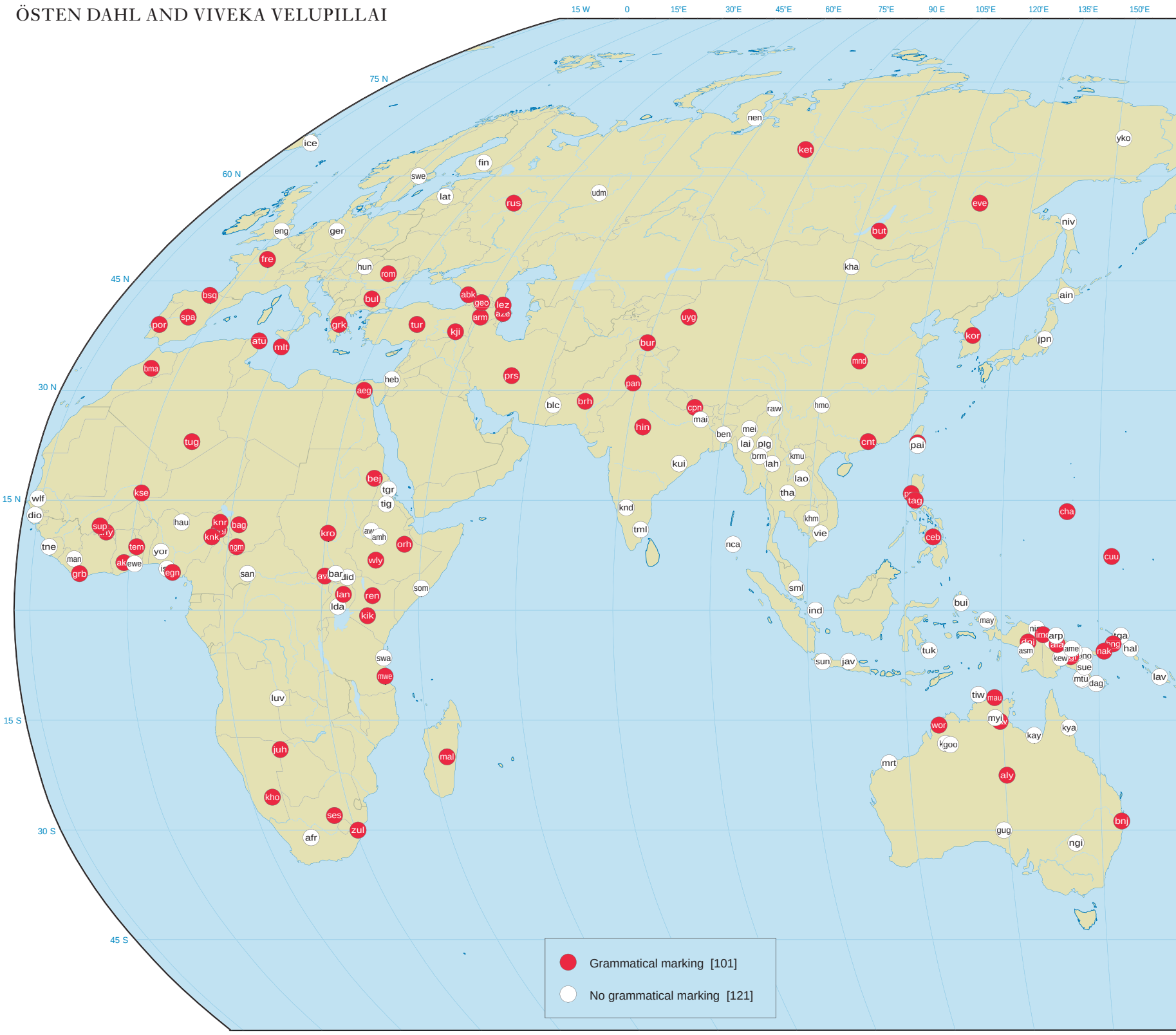
Area where a ‘have’ perfect has developed into a past or a perfective

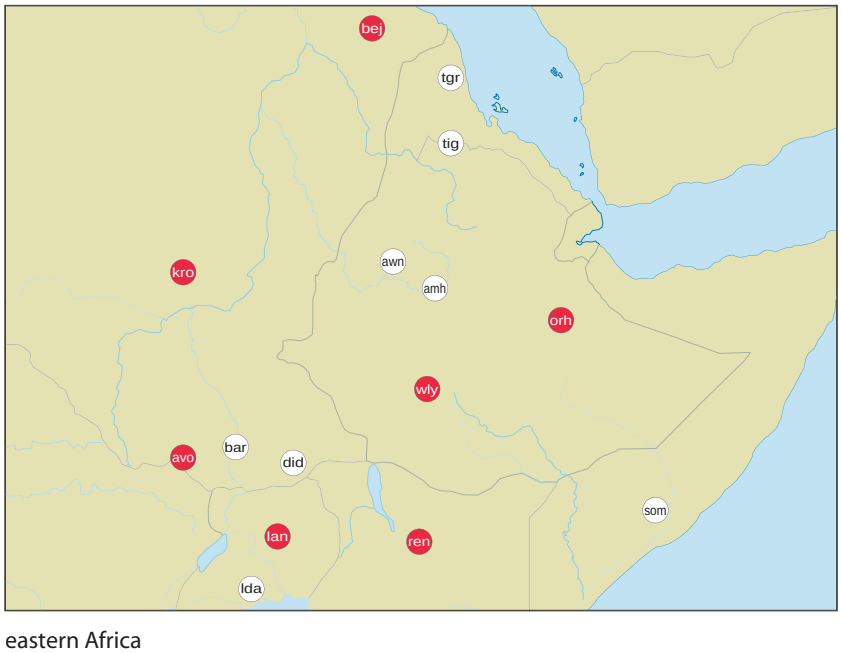
Perfects derived from ‘already’ and ‘finish’—which were noted above as having a special semantics—also show a very marked geographical distribution, being concentrated in South-East Asia and West Africa. (It is possible that more languages should be marked as having this type, but information is sometimes lacking.) It is striking that this coincides with the areas where we find little or no morphological marking of tense and aspect in general, which is reflected on our other maps by the preponderance of white circles.

This page intentionally left blank

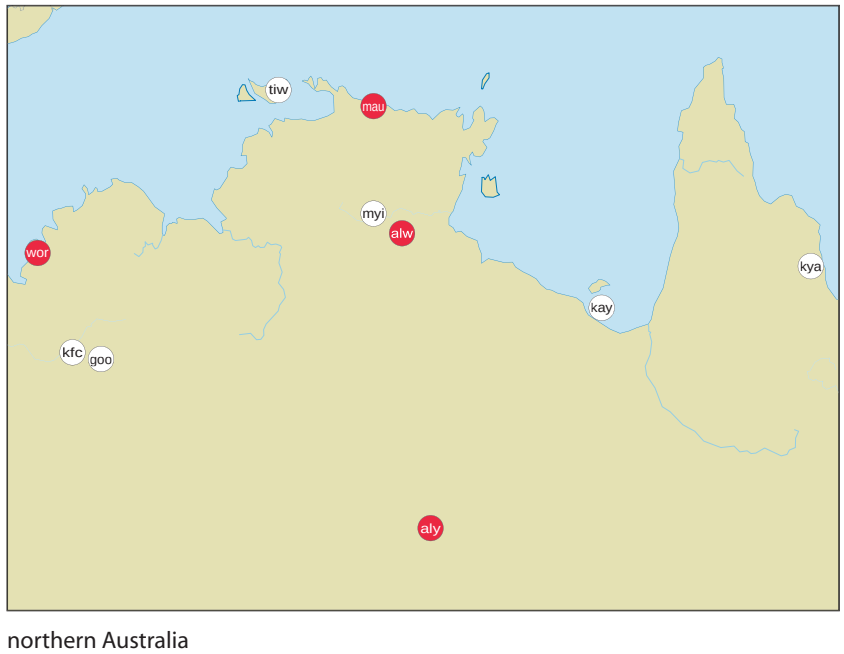
65 Perfective/Imperfective Aspect

ÖSTEN DAHL AND VIVEKA VELUPILLAI

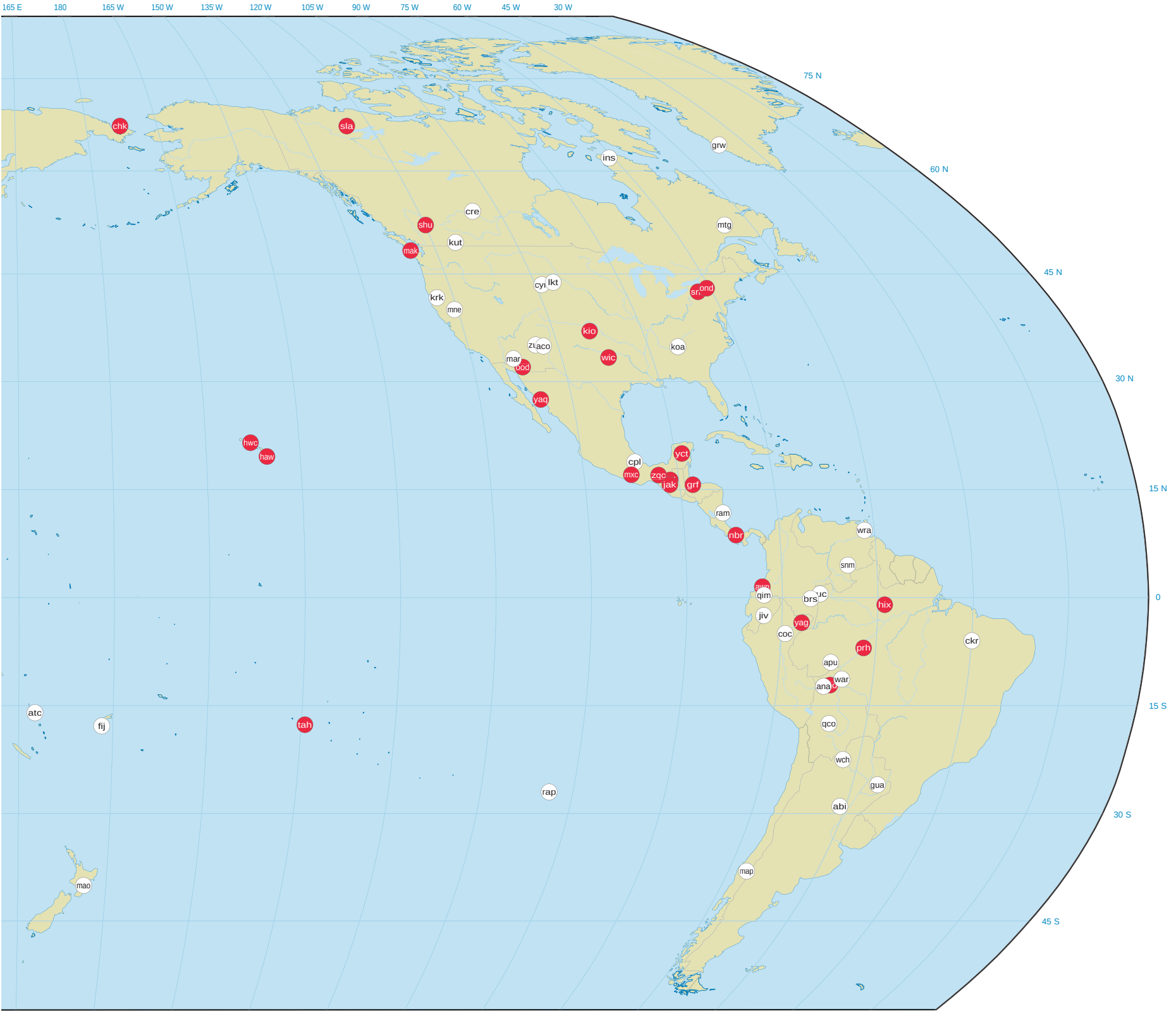


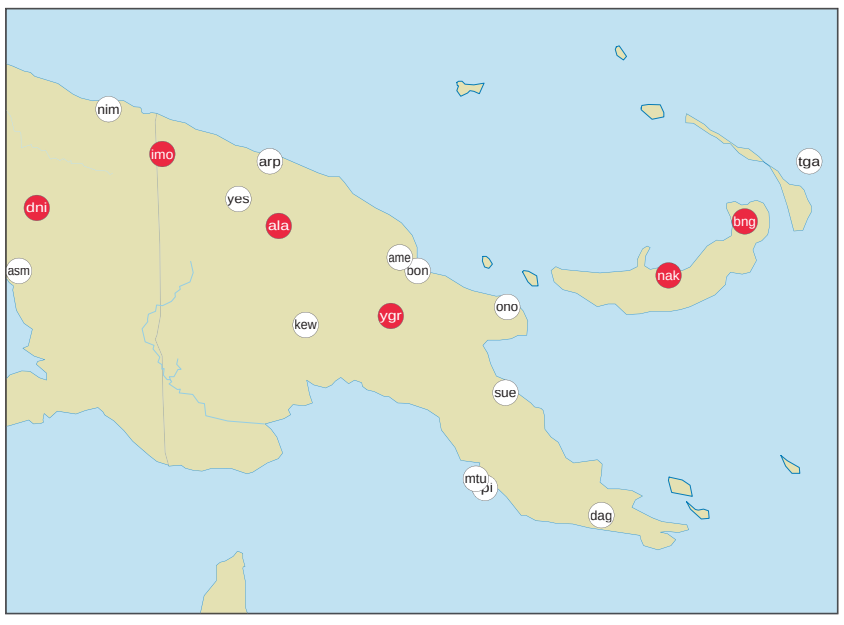


eastern Africa

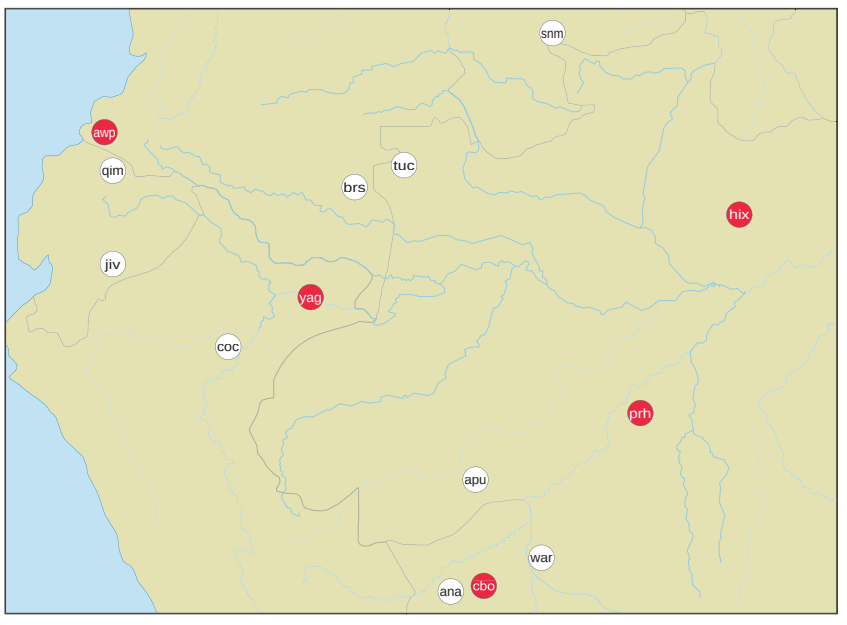


northern Australia





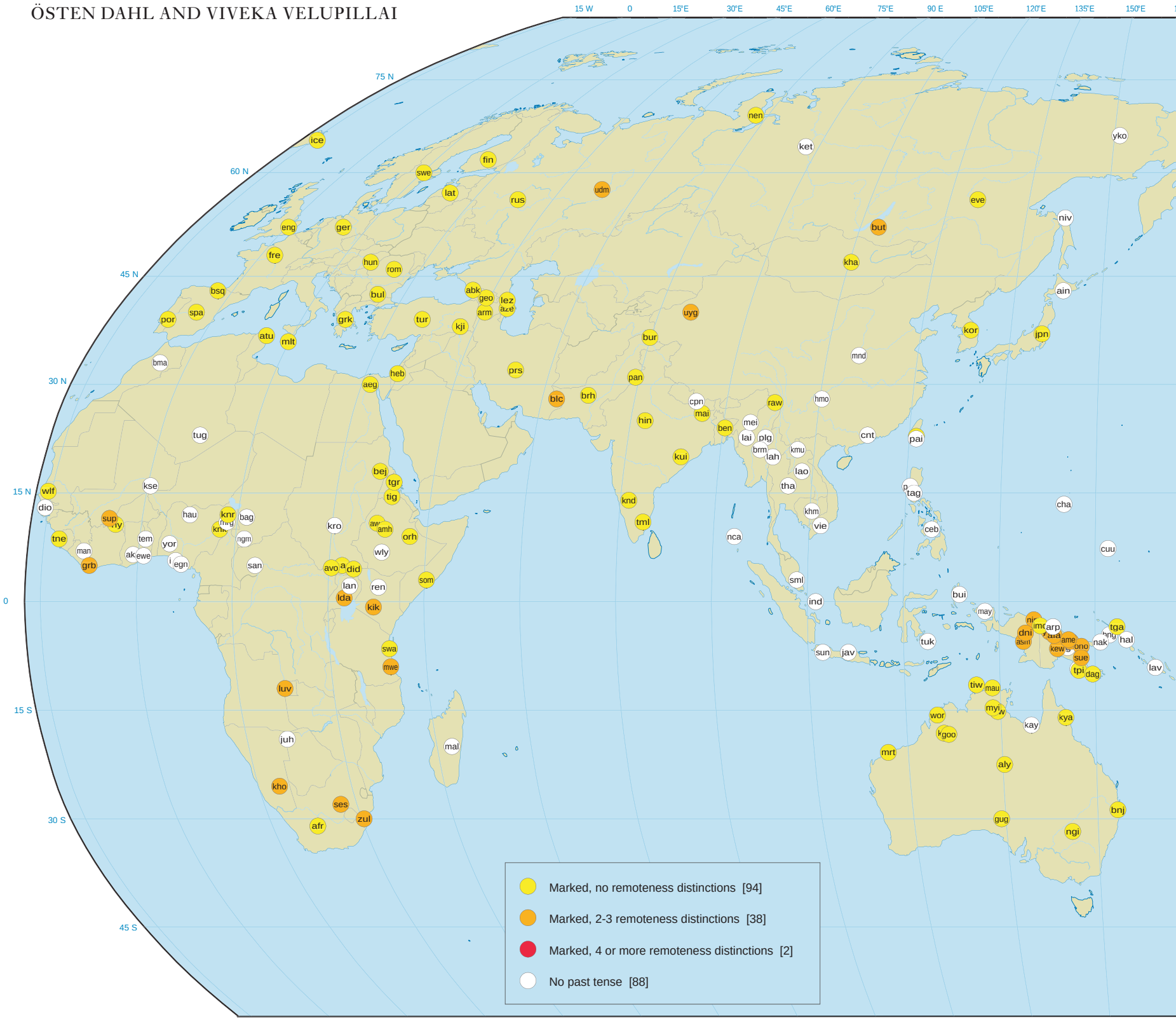
eastern New Guinea

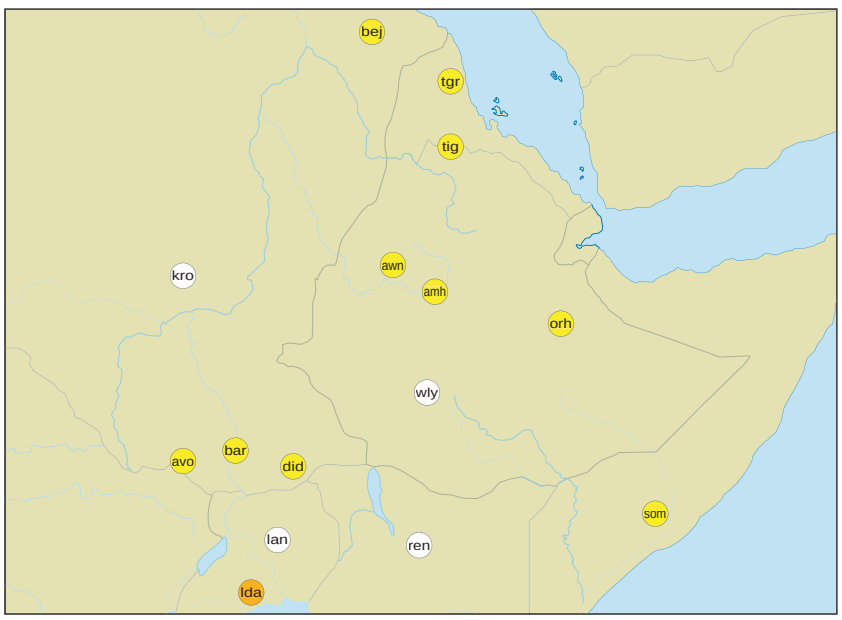


northwestern South America

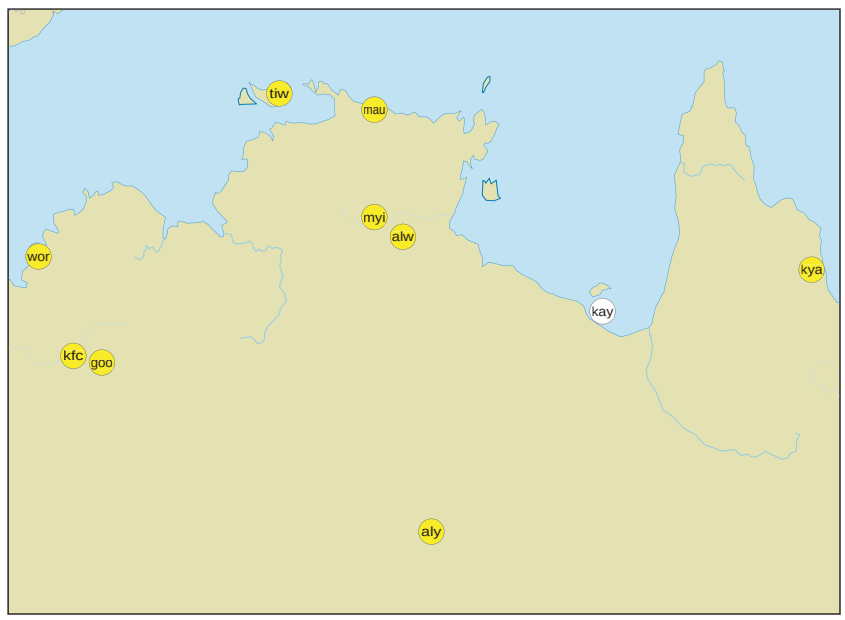
66 The Past Tense

ÖSTEN DAHL AND VIVEKA VELUPILLAI

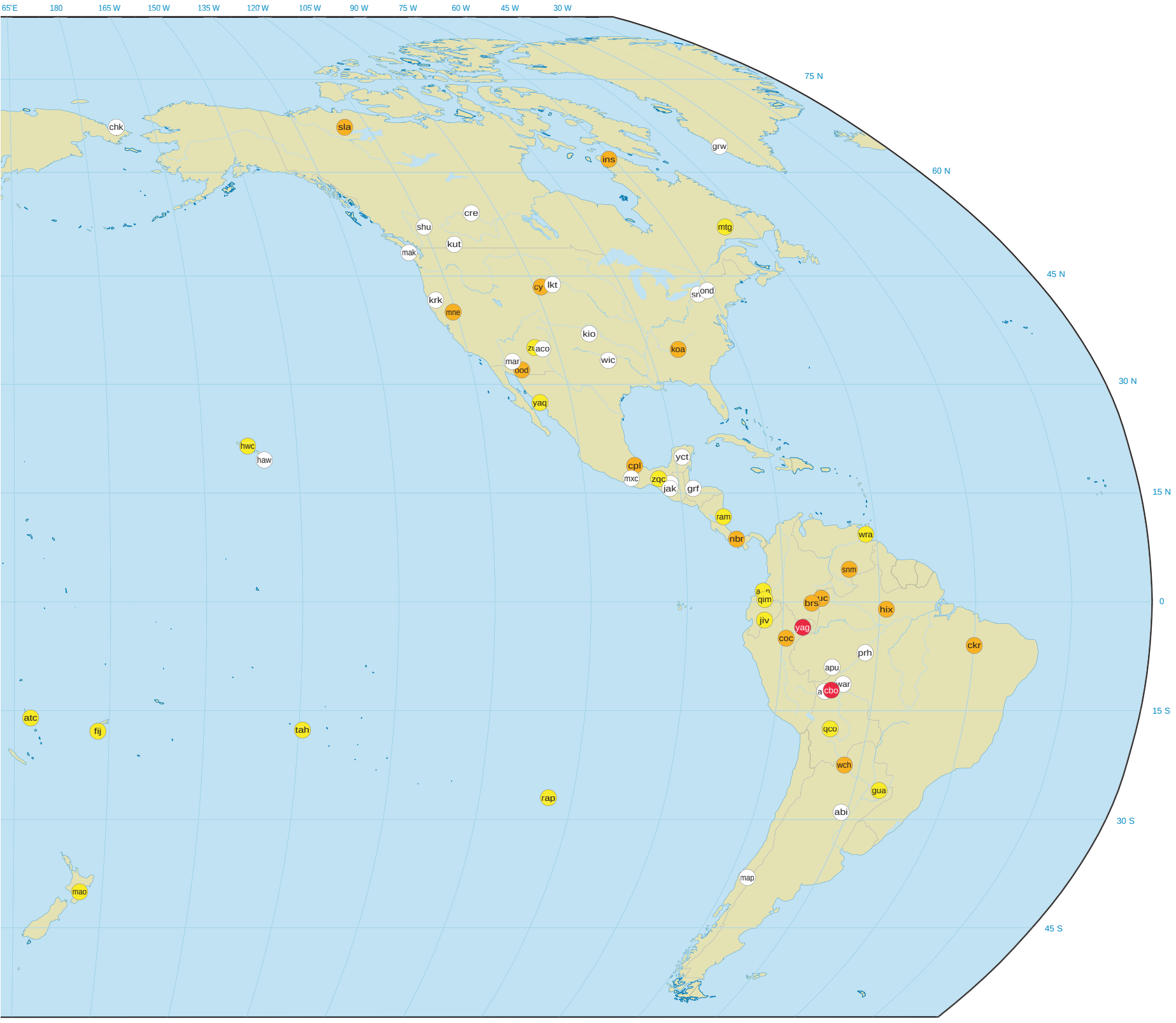


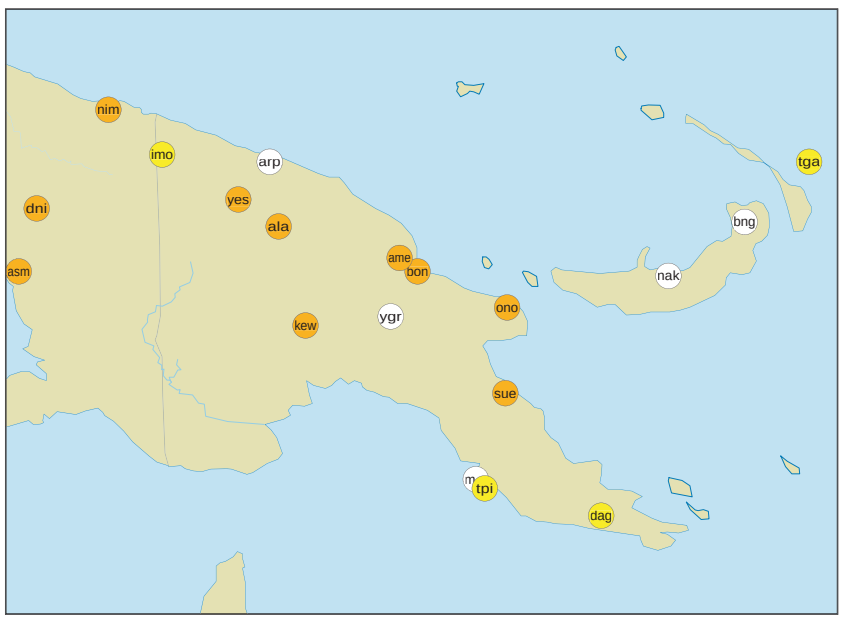


eastern Africa

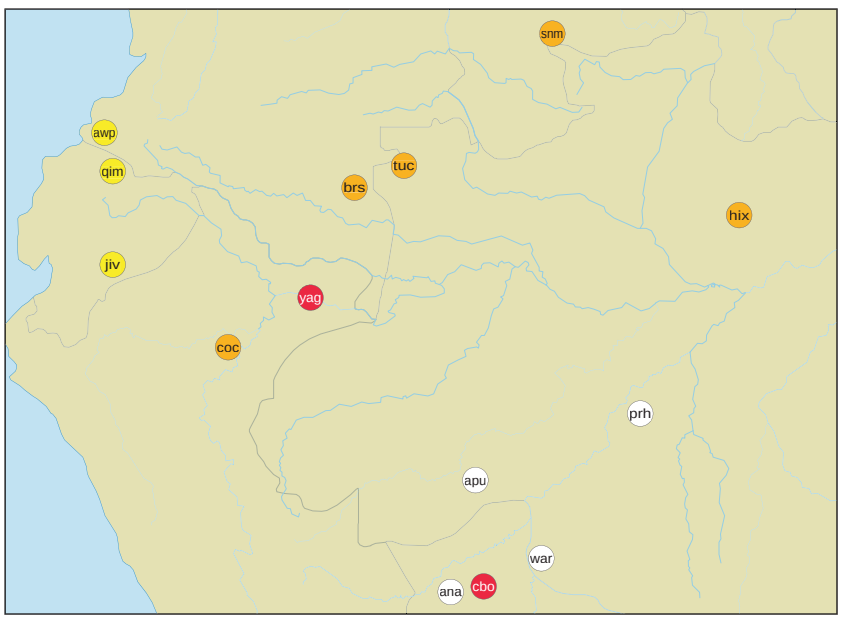


northern Australia





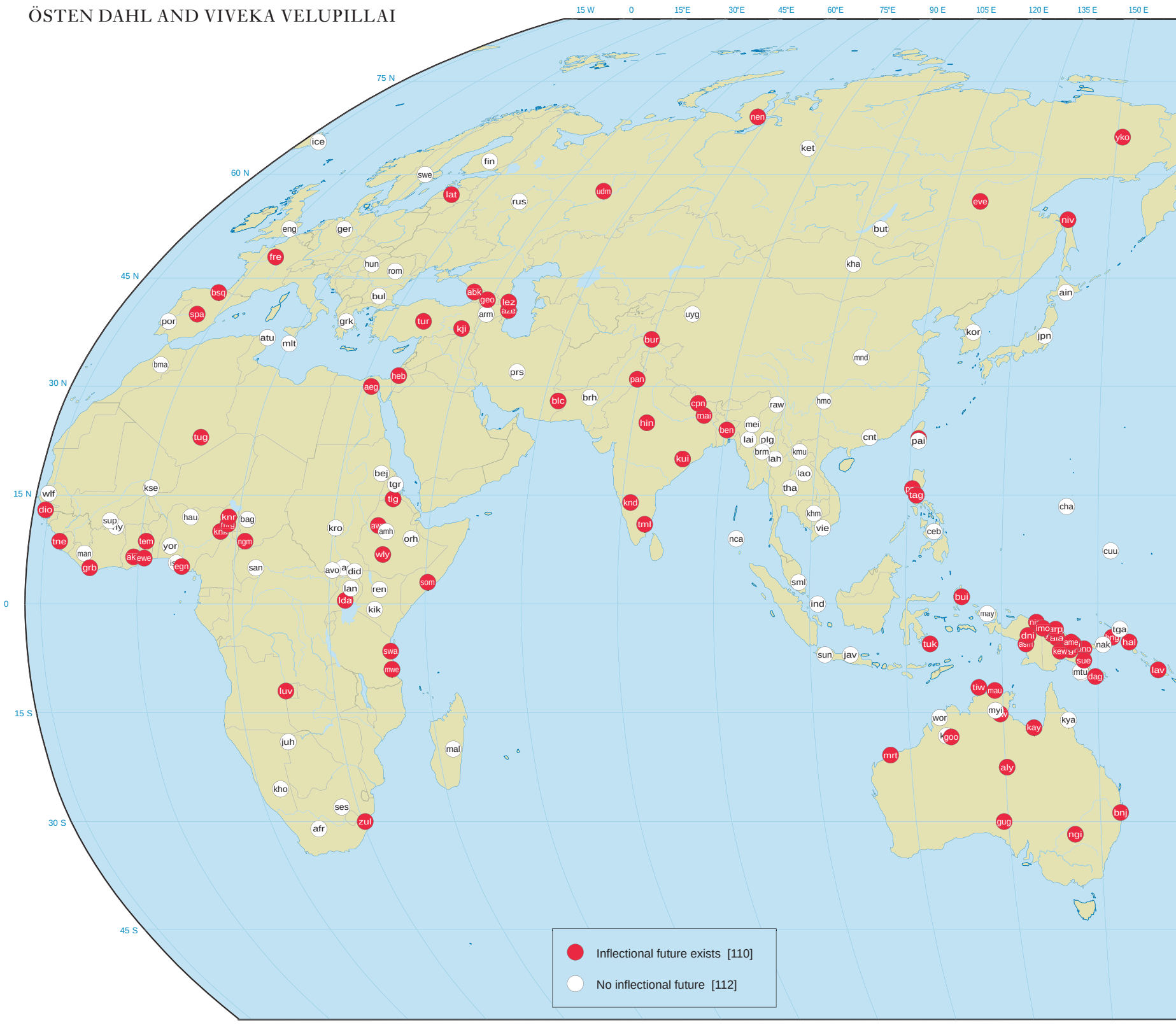
eastern New Guinea

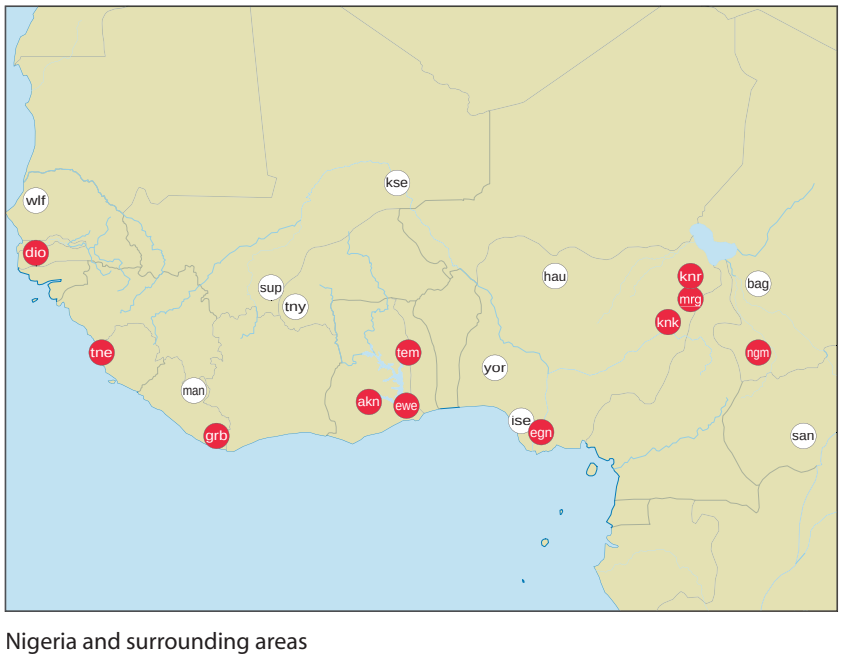


northwestern South America

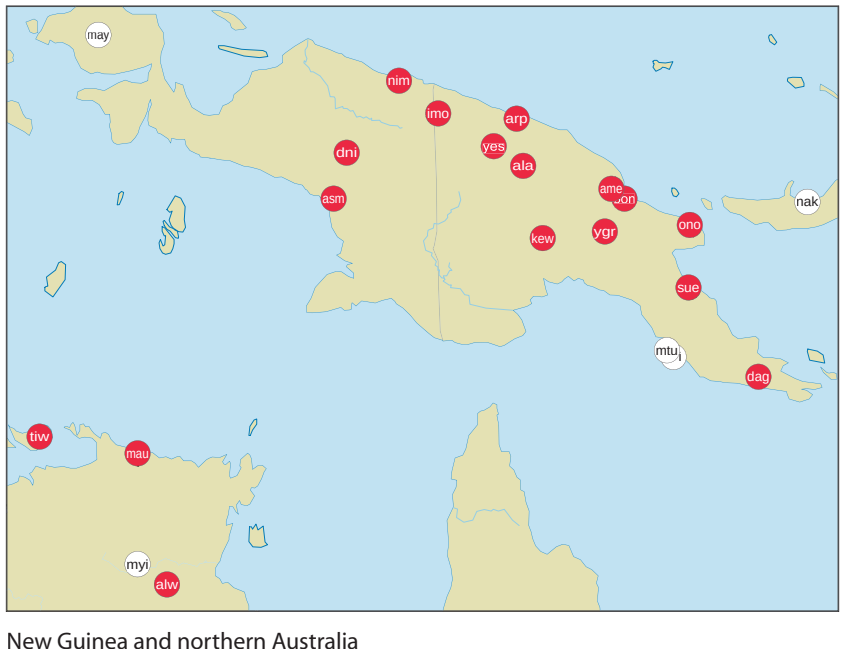
67 The Future Tense

ÖSTEN DAHL AND VIVEKA VELUPILLAI

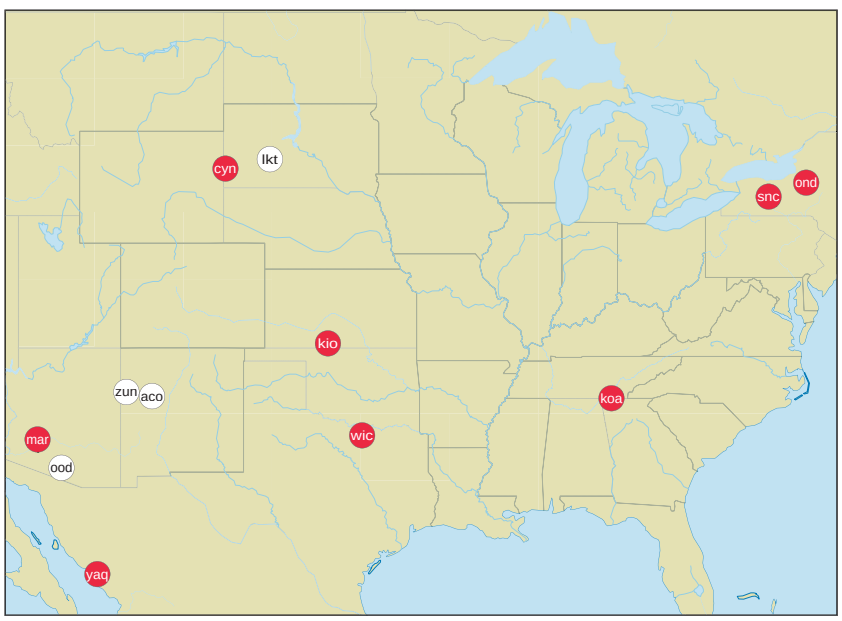




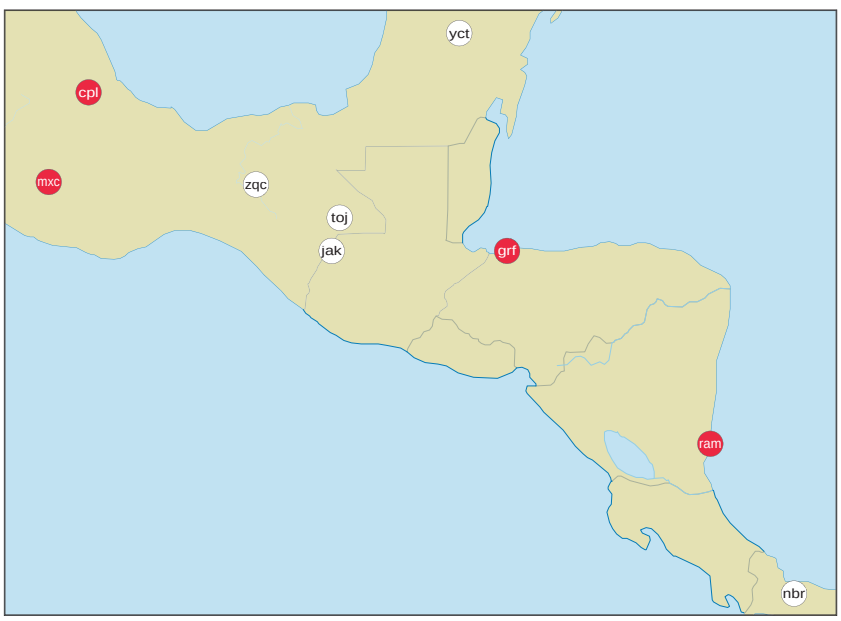
Nigeria and surrounding areas



New Guinea and northern Australia



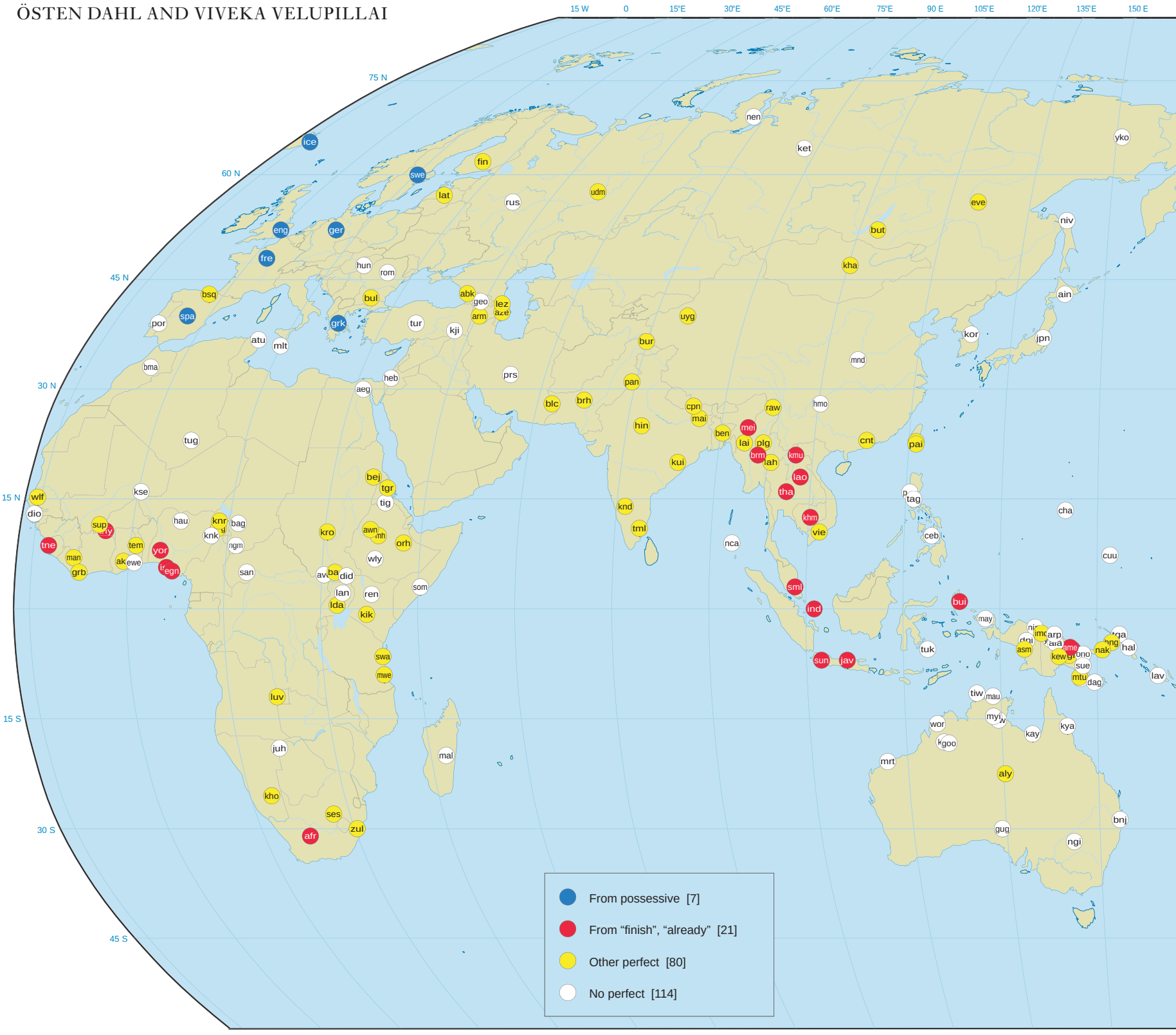
United States

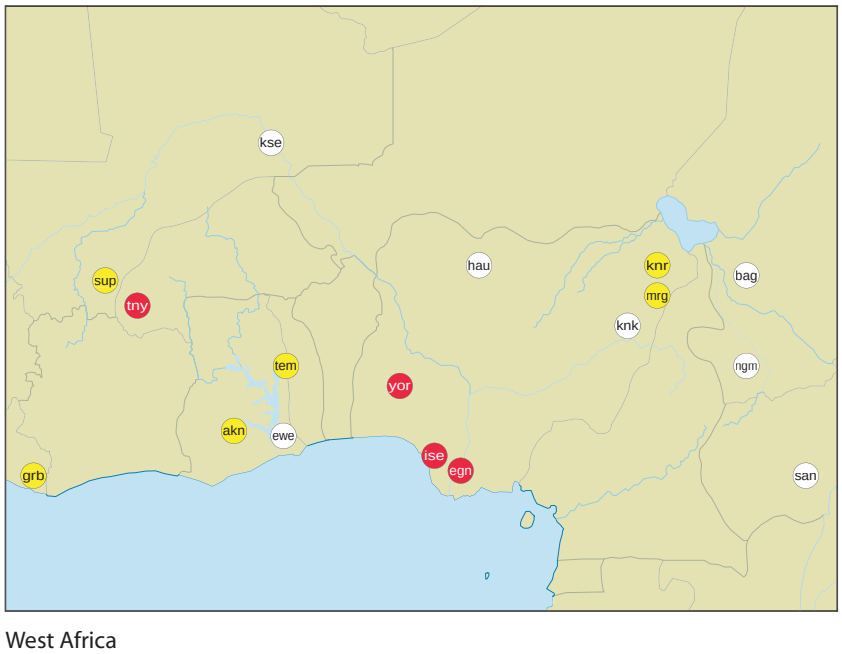


southern Mexico and Central America

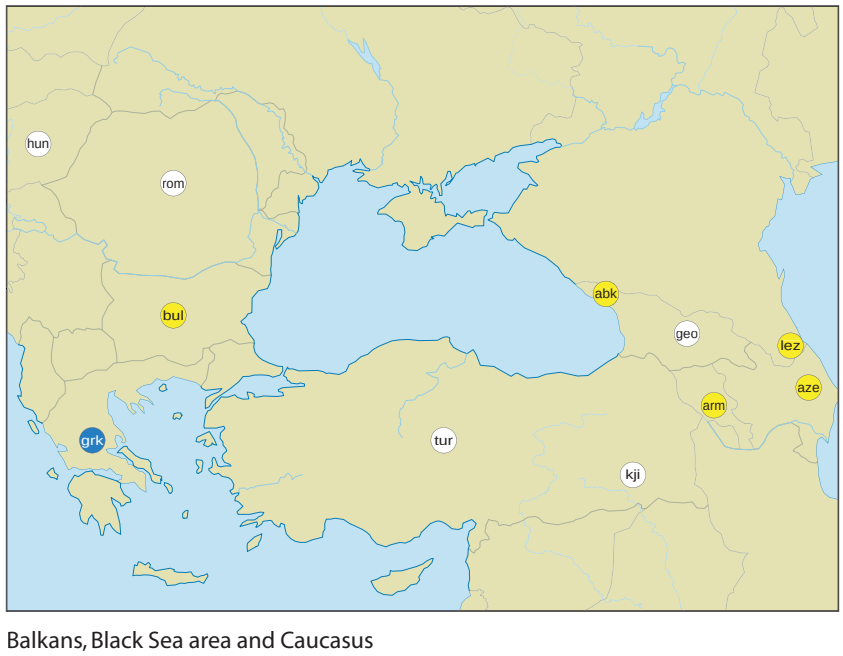
68 The Perfect

ÖSTEN DAHL AND VIVEKA VELUPILLAI

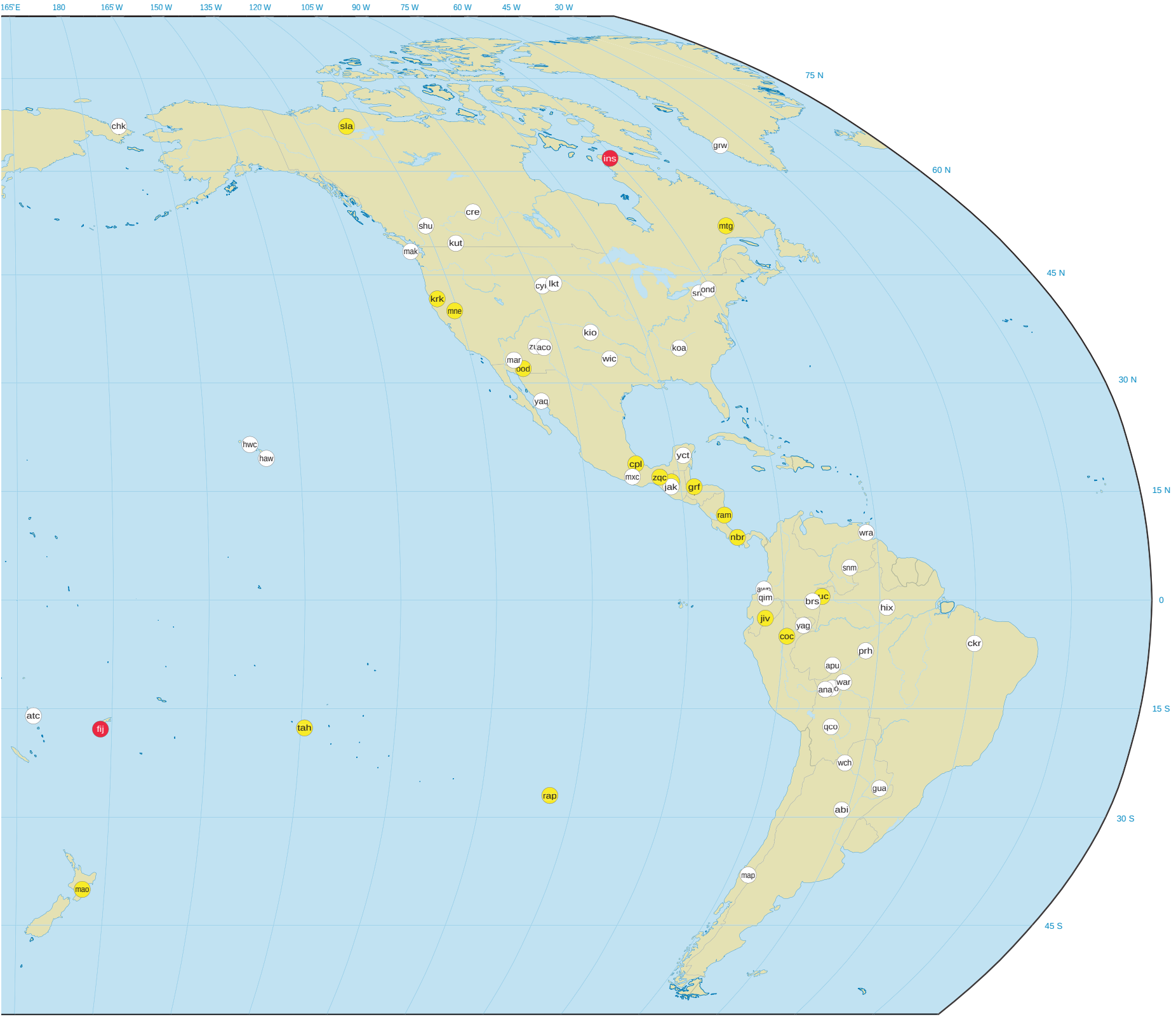


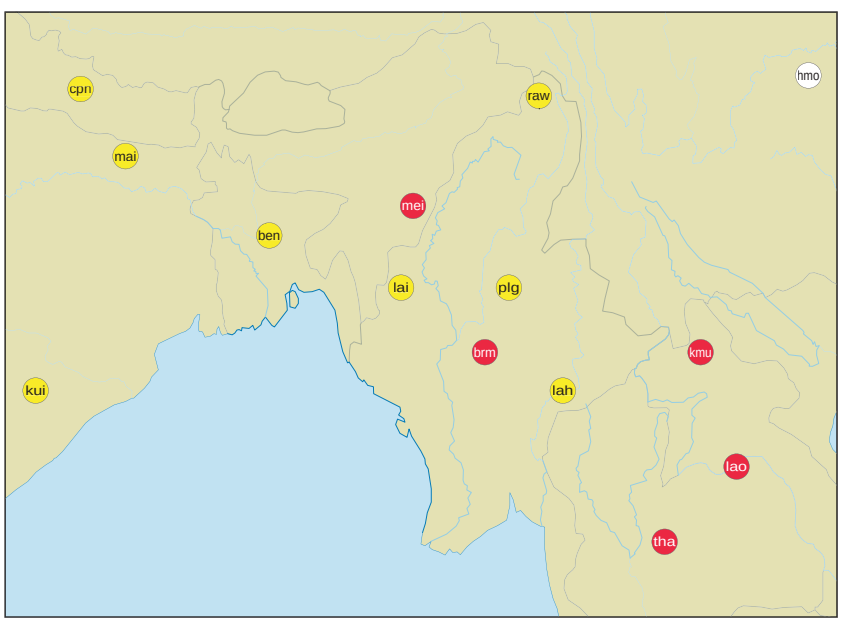


West Africa

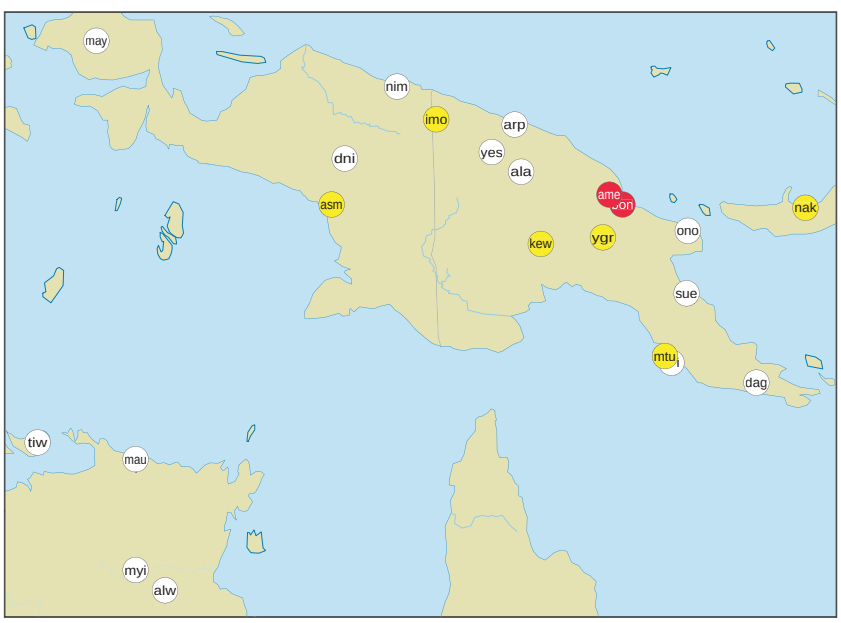


Balkans, Black Sea area and Caucasus





Myanmar and surrounding areas



New Guinea and northern Australia

69 Position of Tense-Aspect Affixes

MATTHEW S. DRYER

1 Defining the values

This map shows the position of tense-aspect affixes. The primary distinction is between **tense-aspect prefixes**, as illustrated by the past tense prefix *ā-* in (1a) from Anywa (Nilotic; Ethiopia and Sudan), and **tense-aspect suffixes**, illustrated by the imperfective suffix *-n* in (1b) from Harar Oromo (Cushitic, Afro-Asiatic; Ethiopia).

(1) a. Anywa (Reh 1996: 199)
Dìmó ā-rúBó kī tìí
Dimo PST-thread.ANTIPASS OBL beads
'Dimo threaded beads [and then . . .]'

b. Harar Oromo (Owens 1985: 100)
sárée-n adíi-n ní iyyi-t-i
dog-NOM white-NOM FOC bark-F-IMPF
'The white dog is barking.'

No attempt is made here to distinguish tense from aspect, and it is frequently difficult to determine from descriptive grammars whether a category ought to be considered tense or aspect. Different descriptions of the same language often differ in whether they characterize a category as one of tense or as one of aspect.

1. Tense-aspect prefixes	150
2. Tense-aspect suffixes	629
3. Tense-aspect tone	11
4. Combination of tense-aspect strategies with none primary	133
5. No tense-aspect inflection	139
total	1062

While the major types of morphological indicators of tense-aspect are prefixes and suffixes, there are three less common morphological ways to indicate tense-aspect: **tone**, **infixes**, and **stem changes**. Tone is the primary method in eleven languages on the map, illustrated in (2) for Lango (Nilotic; Uganda); the initial *a-* in these forms is the first-person singular subject prefix.

(2) Lango (Noonan 1992: 92)
àgíkò 'I stop (something), perfective'
àgíkô 'I stop (something), habitual'
ágìkkò 'I stop (something), progressive'

Apart from the doubling of the second consonant in the progressive form in (2), the aspectual differences are indicated solely by tone. None of the languages examined here appears to use either infixes or stem change as the primary method for tense-aspect, and these types are therefore not shown as types on the map, but, as discussed below, they do combine with other methods in some languages. Changes in the verb stem to express past tense is a common method with irregular verbs in English (e.g. *run, ran; see, saw*), though suffixing (e.g. *walk, walked*) is considered here the primary method for English. If the changes in stems in a language are entirely localized on the initial (or final) segment(s) of the stem, they are not treated here as stem changes, but as prefixes (or suffixes). For example, while the list in (3) of realis and potential (irrealis) forms from Chalcatongo Mixtec (Oto-Manguean; Mexico) does not adhere to a simple pattern, the differences other than tone are all at the beginnings of words, so this language is treated as prefixal on the map.

(3) Chalcatongo Mixtec (Macaulay 1996: 45)

REALIS	POTENTIAL	
ndíso	*kundiso*	'carry'
xátù	*kuxátú*	'be spicy'
žesámá	*kesámá*	'eat'
xíčaʔa	*kačáʔa*	'dance'
xíkó	*kʷíkó*	'spin'
xítú	*kútú*	'work in the fields'
xasú	*kásu*	'close'
kaku	*kákú*	'be born'

Reduplicative processes are treated as prefixes (or as suffixes) if the part of the stem that is reduplicated is less than the whole stem and includes the beginning (or end) of the stem. While complete reduplication is used in some languages for certain aspectual notions like repetitive, continuative, or intensive aspect, as in Turkana (Nilotic; Kenya and Uganda; Dimmendaal 1983: 106), it is not included in the morphological strategies considered here. See Chapter 27 on different types of reduplication.

For many languages, perhaps even a majority, the morphological indicators of tense-aspect on verbs are rather heterogeneous and do not form a single category within the morphological system of the language. For this reason there are many languages which combine two or more of the first three types or which combine one or more of these with infixing or stem changes. If one type is deemed **primary**, either because of the number of relevant morphemes in the language or the apparent frequency in usage, then the language is coded on the map according to that primary method. However, there are many languages which are treated here as lacking a primary method; these are shown on the map as languages with a **combination of tense-aspect strategies with none primary**. The majority of these are languages that employ both prefixes and suffixes. For example, Mesa Grande Diegueño (Yuman; California) has a progressive prefix, illustrated in (4a), and a future suffix, illustrated in (4b).

(4) Mesa Grande Diegueño (Langdon 1970: 147, 159)

a. *ta-ʔ-wa*
PROG-1.SUBJ-sit
'I am/was sitting.'

b. *ʔ-a-x*
1.SUBJ-go-FUT
'I will go.'

There are also many languages which combine prefixes or suffixes with one of the minor strategies. For example, Krongo (Kadugli; Sudan) combines prefixes with tone (Reh 1985: 188). Ocotepec Mixtec (Oto-Manguean; Mexico) combines tone, prefixation, and stem-internal vowel lengthening; compare the following forms of the verb *kunū* 'weave': *kūnū* 'potential', *kúnū* 'continuative', *nkū:nū* 'completive' (Alexander 1988: 250). A few languages combine infixes with some other method: Atayal (Austronesian; Taiwan) employs a past-tense infix and a future prefix (Rau 1992: 47, 50), while Rutul (Daghestanian; Russia) combines infixation with suffixation (Alekseev 1994: 228).

The final type shown on the map is languages with **no tense-aspect affixes** (or stem changes). Such languages will generally code tense and/or aspect categories by means of separate words, either by auxiliary verbs or by noninflecting particles or clitics. For example, in Loniu (Oceanic; Manus Island, Papua New Guinea), tense and aspect are coded by words that precede the verb, illustrated by the present tense and stative aspect words in (5).

(5) Loniu (Hamel 1994: 107)
yo leʔi tɔ ehe
1SG PRES STAT lie.down
'I am lying down.'

Languages with tense-aspect affixes vary as to the centrality of these affixes to the languages. In some languages, verbs obligatorily inflect

for tense or aspect, although one of the categories may be realized by the absence of any overt tense-aspect morphology. In Lango, for example, illustrated above in (2), each verb form must be in one of the three aspects shown there. In other languages, the tense-aspect affixes play a minor role in the system. In Indonesian, tense-aspect is generally expressed by separate words, as in (6a), but there is one suffix that occurs on a number of transitive verbs to indicate repetitiveness or thoroughness, as in (6b).

(6) Indonesian (Sneddon 1996: 199, 94)

a. *Kami akan makan nanti.*
1PL FUT eat soon
'We will eat soon.'

b. *Dia men-cium-i pacar-nya.*
3SG ACTIVE-kiss-REPET girlfriend-3SG.POSS
'He kissed his girlfriend repeatedly.'

The map does not distinguish languages in which tense-aspect affixes are central from those in which they are less so; because of the single suffix illustrated in (6b), Indonesian is shown on the map as having tense-aspect suffixes.

In deciding whether a morpheme indicating tense-aspect is an affix or a separate word, I follow the orthography of my sources. In some instances, more careful linguistic analysis would probably lead to an alternative analysis. Analyses are often strongly influenced by how morphemes have been analysed in related languages or languages in the same region. In some cases, the facts of a language may be somewhat indeterminate as to whether a particular morpheme should be treated as an affix or a separate word.

Apart from familiar tense categories like past, present, and future, the affixes shown on the map code aspectual categories which are characterized in grammatical descriptions with labels such as perfective, imperfective, progressive, continuative, repetitive, and habitual. In some languages, the indication of tense and aspect combines with pronominal affixes, especially for subject, so that the affixes in question are often described as agreement affixes or subject affixes which vary for tense or aspect. Such affixes which code some other category in addition to tense-aspect are included here as tense-aspect affixes, as long as one of the grammatical features they code is tense or aspect.

Many languages make a fundamental distinction in their verbal morphology that authors of grammars describe by means of the labels **realis** and **irrealis**. If the irrealis forms are obligatory for referring to the future, and are normally the sole means of indicating future, then the marking of realis versus irrealis is treated here as an instance of tense-aspect marking. In Tukang Besi (Austronesian; Indonesia), for example, there are two sets of subject prefixes, one realis set and one irrealis set, as in (7).

(7) Tukang Besi (Donohue 1999: 153)

a. ***no**-baiara-'e*
3.SUBJ.REALIS-pay-3.OBJ
'She has paid it.'

b. ***na**-baiara-'e*
3.SUBJ.IRREALIS-pay-3.OBJ
'She is going to pay it.'

Because the contrast in the simplest contexts is that of nonfuture versus future, these prefixes are treated as tense-aspect prefixes for the purposes of this map.

2 Geographical distribution

Of the two major types, prefixes and suffixes, it is clear from the map that tense-aspect suffixes are overwhelmingly more common than tense-aspect prefixes. There are a number of areas in which prefixes are extremely rare. In the entire mainland of Europe and Asia, there are only three languages shown on the map as employing prefixes as the primary strategy: Ket (Yeniseian; Russia; Werner 1997: 154–5), Jiarong (Tibeto-Burman; China; Lin 1993: 231), and Temiar (Aslian, Mon-Khmer; Malaysia; Benjamin 1976: 169), and the last of these is well down the Malay Peninsula and barely on the mainland. Only two languages in South America and only three in Australia are shown as employing prefixes as the primary strategy. And while the map shows a number of languages in New Guinea with tense-aspect prefixes, most of these languages are Austronesian; the non-Austronesian languages of New Guinea overwhelmingly employ suffixing for tense-aspect.

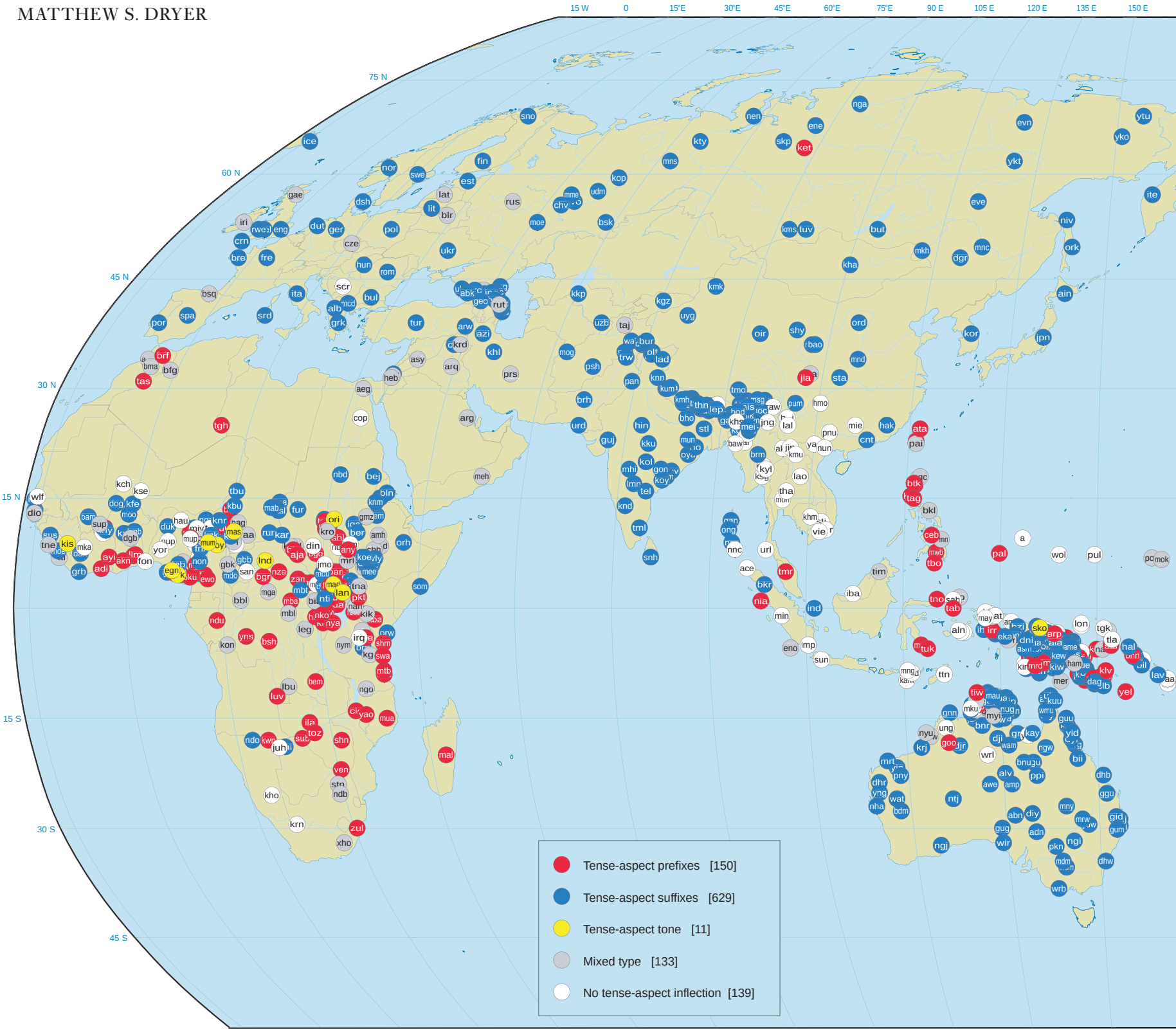
The strongest exception to the preference for tense-aspect suffixes is found among Austronesian languages; although many lack tense-aspect affixes, among those that do have them, they are more generally prefixes. There are two areas with tense-aspect prefixes among Austronesian languages: (i) a strip of languages extending from Taiwan south through the Philippines to Sulawesi in eastern Indonesia; and (ii) Austronesian languages of eastern New Guinea, some on the mainland, some on islands like New Britain. Another area in which tense-aspect prefixes are common is Africa. While there are approximately as many languages with tense-aspect suffixes in Africa as there are with prefixes, languages with prefixes are distributed over much of Africa and are found in a number of different subgroups within Niger-Congo (including Bantu, Kwa, Adamawa-Ubangian, Kordofanian) as well as many subgroups within Nilo-Saharan and a few Afro-Asiatic languages within Chadic and Berber. Finally, there are two areas in North America where tense-aspect prefixes are common: (i) Mesoamerica; and (ii) Canada and the northern United States, represented by Athapaskan and Algonquian languages. There are a few languages in the southwest United States and adjacent areas in north-west Mexico with tense-aspect prefixes, but two of these are Athapaskan languages, closely related to languages in northern Canada.

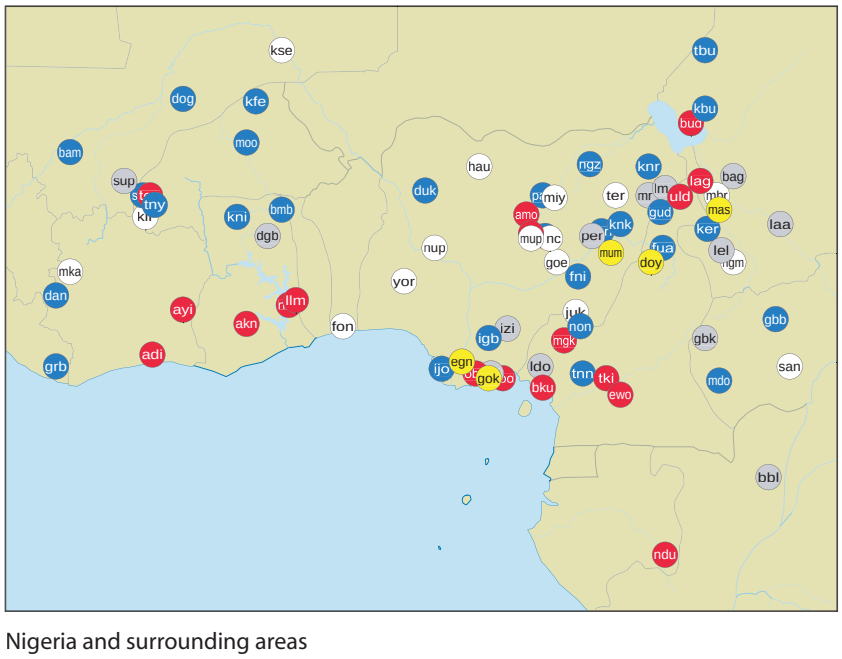
With one exception, the languages employing tone as the primary method for coding tense-aspect are restricted to Africa, but they are spoken over a wide area stretching from Guinea in West Africa to Uganda in East Africa. The one language not in Africa is Sko (Sko family; Papua, Indonesia; Voorhoeve 1971: 57). There are many other languages in which tone combines with some other method. A number of these are Mixtecan or Chinantecan languages spoken in Mesoamerica.

Languages lacking tense-aspect affixes are most common in two areas in which isolating languages, ones with little or no inflectional morphology, are common: (i) South-East Asia and Austronesian languages of Indonesia and the Pacific; (ii) central Africa, especially around Nigeria. On this map, the first of these two areas extends northwestward to include many Tibeto-Burman languages which lack tense-aspect affixes but have other inflectional morphology. It should be noted that there are also polysynthetic languages without tense-aspect affixes. For example, Kutenai (isolate; western North America) expresses all tense and aspect notions by preverbal particles but is otherwise polysynthetic, and the neighbouring Salishan language Shuswap is similar (Kuipers 1974: 45, 74, 80–1).

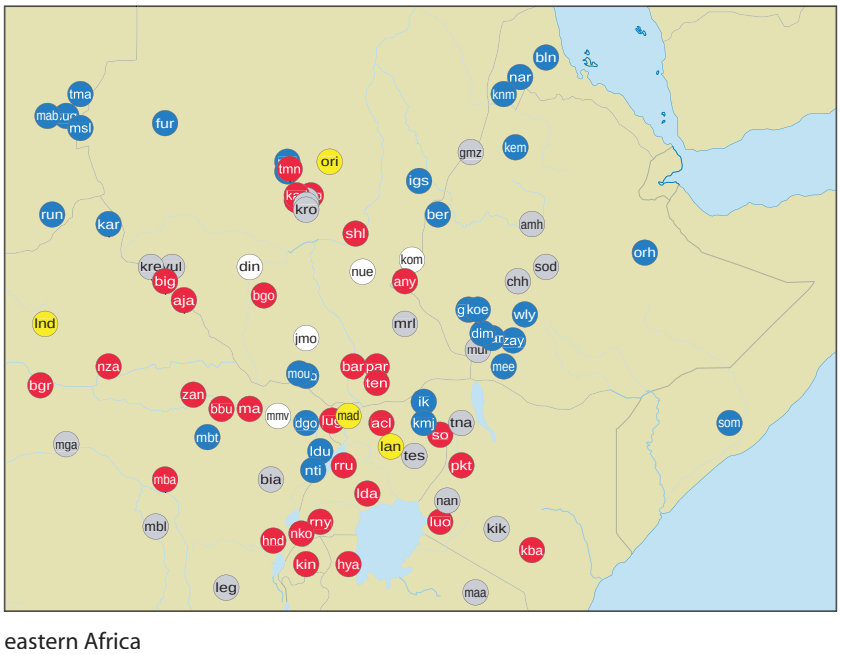
69 Position of Tense-Aspect Affixes

MATTHEW S. DRYER

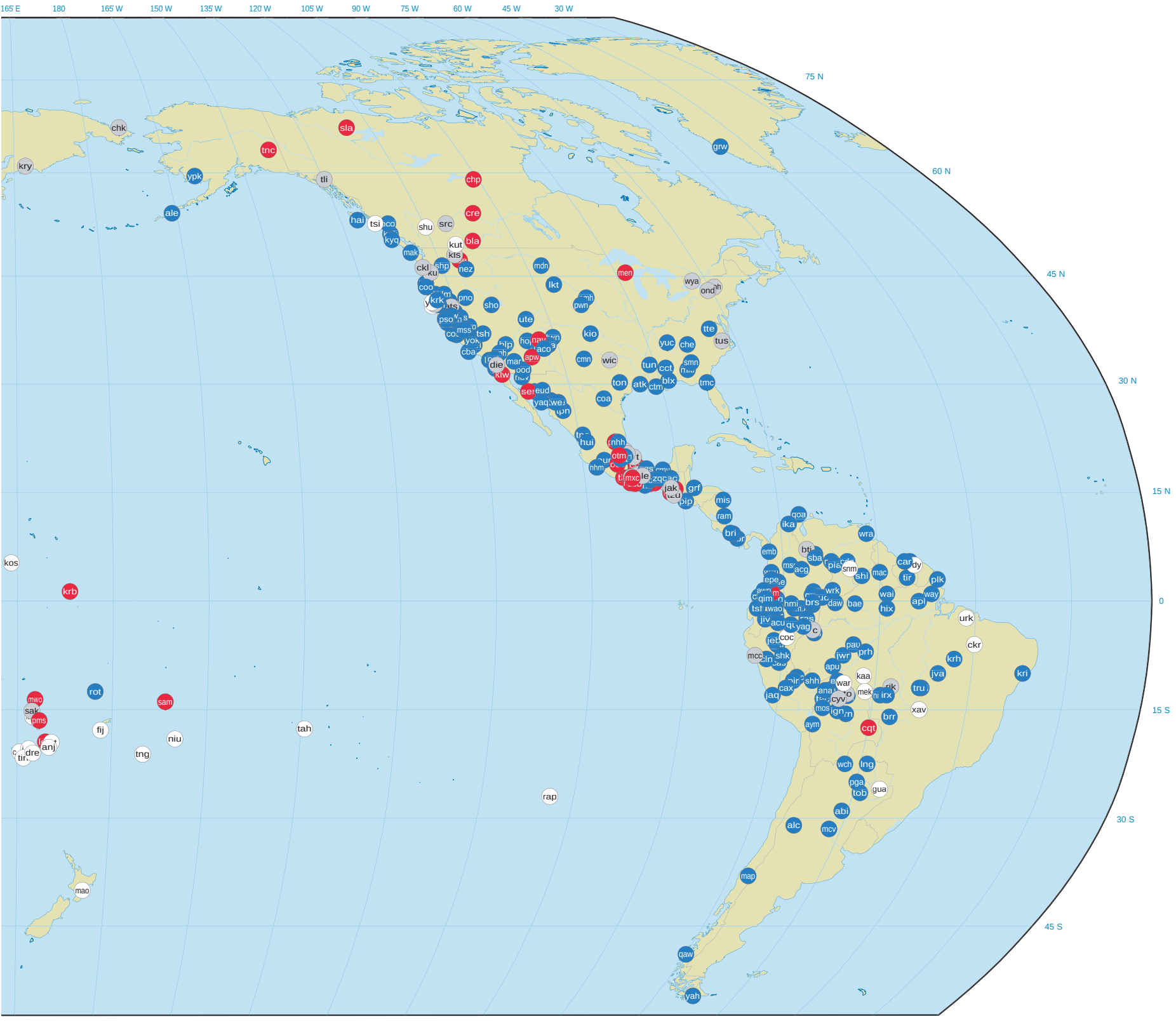


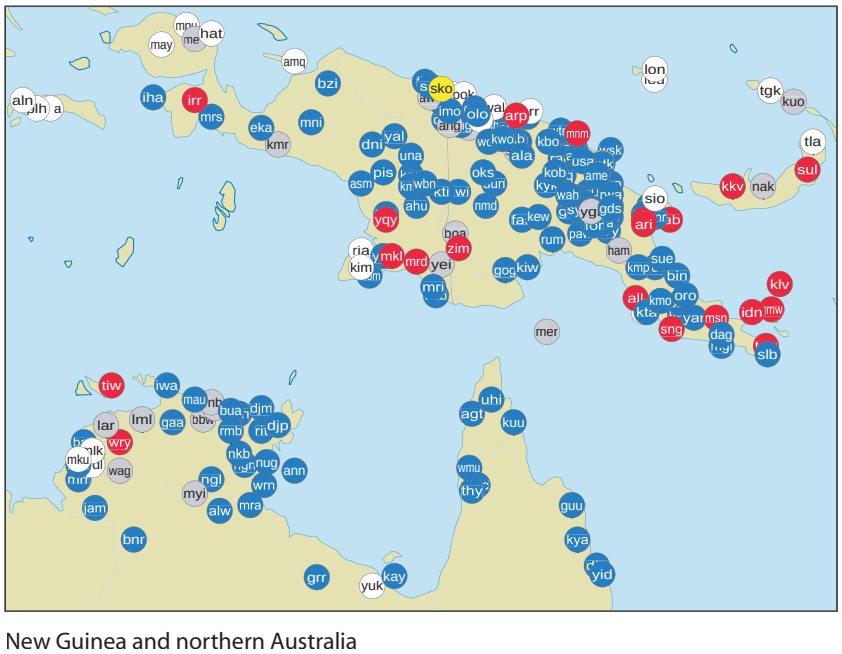


Nigeria and surrounding areas

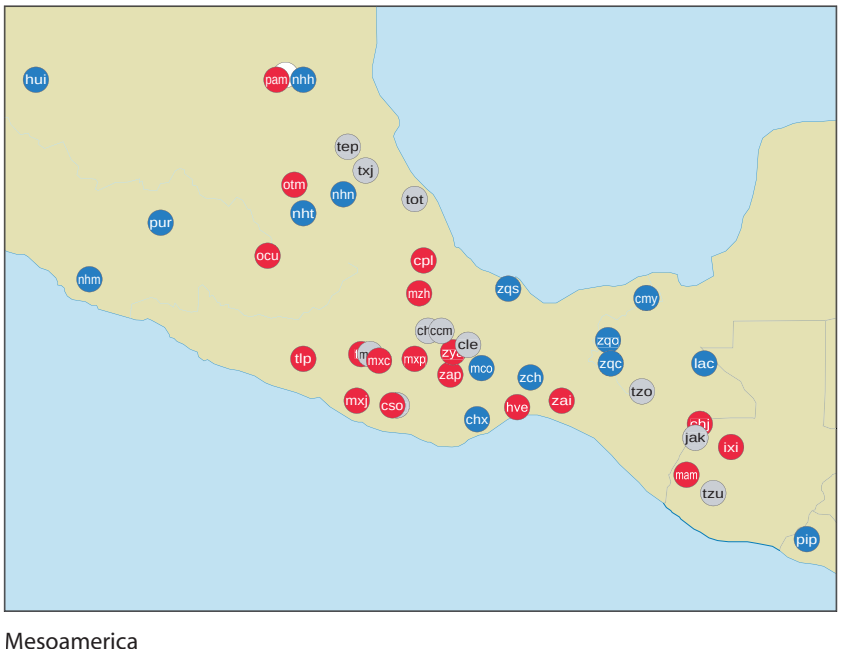


eastern Africa





New Guinea and northern Australia



Mesoamerica

70 The Morphological Imperative

JOHAN VAN DER AUWERA AND LUDO LEJEUNE
(with Umarani Pappuswamy and Valentin Goussev)

1 Definition of values

This map shows to what extent languages have second-person singular and plural imperatives as dedicated morphological categories.

1.	The language has morphologically dedicated second singular as well as second plural imperatives.	292
2.	The language has morphologically dedicated second singular imperatives but no morphologically dedicated second plural imperatives.	42
3.	The language has morphologically dedicated second plural imperatives but no morphologically dedicated second singular imperatives.	2
4.	The language has morphologically dedicated second-person imperatives that do not distinguish between singular and plural.	89
5.	The language has no morphologically dedicated second-person imperatives at all.	122
	total	547

The first type shown on the map includes languages that have **special morphological marking both for the imperative second singular and for the imperative second plural**, as in example (1) from Limbu (Tibeto-Burman; Nepal).

(1) Limbu (van Driem 1987: 188)

a. *Ips-Ø-ɛʔ!*
sleep-2SG-IMP
'Sleep!'

b. *Ips-amm-ɛʔ!*
sleep-2PL-IMP
'Sleep!'

As (1a) illustrates, zero markers count as markers, too.

The second type shown is that of languages that possess **morphologically specialized imperative forms for the second singular but not for the second plural**. In Lingala (Bantu; Democratic Republic of Congo), the imperative only exists in the second singular. To address an order to more than one addressee, speakers have to resort to another form, for instance, a subjunctive form.

(2) Lingala (Meeuwis 1998: 28)

a. *Sál-á!*
work-EPV.IMP.SG
'Work!'

b. ***Bó**-sál-a!*
SBJV.2PL-work-EPV
'Work!'

An alternative analysis would say that Lingala shows syncretism between its imperative and subjunctive in the second-person plural. For the purposes of this chapter, however, the debate is not relevant. Either way, it is correct to say that Lingala lacks a specifically imperative second plural. Ceteris paribus, this remark is relevant to languages of type three and four also.

The third type shows languages with **dedicated morphology for the second plural imperative, but not for the second singular imperative**, as in Latvian. To address an order to a single addressee, Latvian speakers use the morphological indicative present.

(3) Latvian (Holst 2001: 139)

a. *Runā!*
speak.IND.PRES.2SG
'Speak!'

b. *Runājāt!*
speak.IMP.2PL
'Speak!'

Languages of the fourth type have a **specialized second-person imperative that does not distinguish between singular and plural**, as in Paiwan (Austronesian; Taiwan)

(4) Paiwan (Egli 1990: 282)
*Kan-**u**!*
eat-IMP.2
'Eat!'

The final type shows languages that have **no morphologically dedicated second-person imperative at all, neither for the singular nor for the plural**. In Nunggubuyu (Australian, Gunwinyguan; Northern Territory, Australia) the forms that are typically used for expressing orders are morphological future indicatives.

(5) Nunggubuyu (Heath 1984: 339, 343, 344, 348)

a. *Ba-bura:!*
2SG-sit.FUT.IND.CONT
'Sit!' *or* 'You will sit.'

b. *Numburu-bura:!*
2PL-sit.FUT.IND.CONT
'Sit!' *or* 'You will sit.'

In some languages the presence or absence of dedicated imperative morphology is dependent on the type of verb. A language will be classified as having specialized imperative morphology only if this morphology is available for a large set of verbs. Languages often have small sets of irregular verbs; if the imperatives of these verbs are also irregular, their properties will not be taken into account. In Georgian, for instance, most imperatives have the same form as indicative aorists or subjunctive presents (see 6a–b). However, the verb 'come' has a specialized imperative form (see 6c).

(6) Georgian (Vogt 1971: 196, 198)

a. *Dac'ere*.
write.IND.AOR.2SG
'You wrote.' *or* 'Write!'

b. . . . *icode* *rom* . . .
know.SBJV.PRES.2SG that
'. . . that you know that . . .' *or* 'Know that . . .'

c. *Modi!*
come.IMP.2SG
'Come!'

Thus, Georgian is classified as belonging to the fifth type (no morphologically dedicated imperative at all).

Languages may have more than one imperative paradigm, distinguished along parameters like tense (most typically present versus future), aspect (e.g. perfective versus imperfective), politeness, movement towards or away from the speaker, voice, or transitivity. If these strategies differ with respect to the morphological dedication of the second singular and plural, we have only taken the most general type into account, e.g. the present imperative rather than the future one. In the few cases in which both would seem equally general and only one shows morphological dedication, we have coded the

language on the basis of just one of the two strategies. An illustration comes from Nkore-Kiga (Bantu; Uganda), in which the intransitive second singular imperative is morphologically dedicated while the transitive one, comprising an object clitic, is not, resorting to a subjunctive strategy and changing the stem-final *-a* into *-e*.

(7) Nkore-Kiga (Taylor 1985: 163, 12)

a. *Kora!*
work.IMP.2SG
'Work!'

b. *Ki-reete!*
it-bring.SBJV
'Bring it!'

In many languages the second-person imperative, especially the imperative second singular, employs a bare stem. This form may be employed elsewhere in the verbal paradigm. Dutch has a second-person imperative that does not distinguish between singular and plural. This same form is also used as a first-person indicative present.

(8) Dutch

a. *Zing!*
sing.IMP.2
'Sing!'

b. *Ik zing.*
I sing.IND.PRES.1SG
'I sing.'

Since there is no plausible semantic connection between a second-person imperative and a first-person indicative present, we consider Dutch as having dedicated imperative morphology.

In Ingush (Nakh-Daghestanian; northern Caucasus) most verbs have an imperative that is homophonous with the infinitive.

(9) Ingush (Nichols 1994: 106)
Dieš-a!
read-IMP *or* read-INF
'Read!' *or* 'to read'

The syncretism would allow one to say that the form in (9) is not a morphological imperative at all, but only an infinitive, which, when occurring by itself, serves an imperative function (see Guérin 1998: 261). It is furthermore uncontroversial that languages do indeed employ infinitives to express commands—see Dutch (10).

(10) Dutch
Zingen!
sing.INF
'Sing!'

Under this analysis, Ingush could be taken to lack a simple morphological second-person imperative. However, for the purposes of this chapter, we have followed the rule that if the form used for a second-person imperative is identical to another second-person form elsewhere in the verbal paradigm, with the same number, then we do not speak of morphological dedication. Thus the Latvian second singular imperative was not considered morphologically dedicated. If, however, the form used for a second-person imperative is identical to a form other than a same-number second person, then we do speak of morphological dedication. Thus Dutch *zing* and Ingush *dieša* will be considered to be morphologically dedicated.

2 Geographical distribution

The most common strategy worldwide is to morphologically dedicate both the imperative second singular and plural. The only area where morphological dedication is absent is South-East Asia; the reason is that the languages in this area are overall lacking in morphology. Another area with a good many morphologically nondedicated imperatives is sub-Saharan Africa. The lack of morphological dedication may be typical for Khoisan languages; in the equatorial area it is found in various Niger-Congo families, perhaps most often in the Adamawa-Ubangi languages. It is found in the Americas, New Guinea, and northern Australia. In Eurasia it is vary rare, and in Europe it is found only in English and Hungarian. To some extent, the absence of a morphological imperative correlates with the language's general lack of morphology, e.g. English, Trumai (isolate; Brazil), or Ewe (Kwa, Niger-Congo; Togo and Ghana).

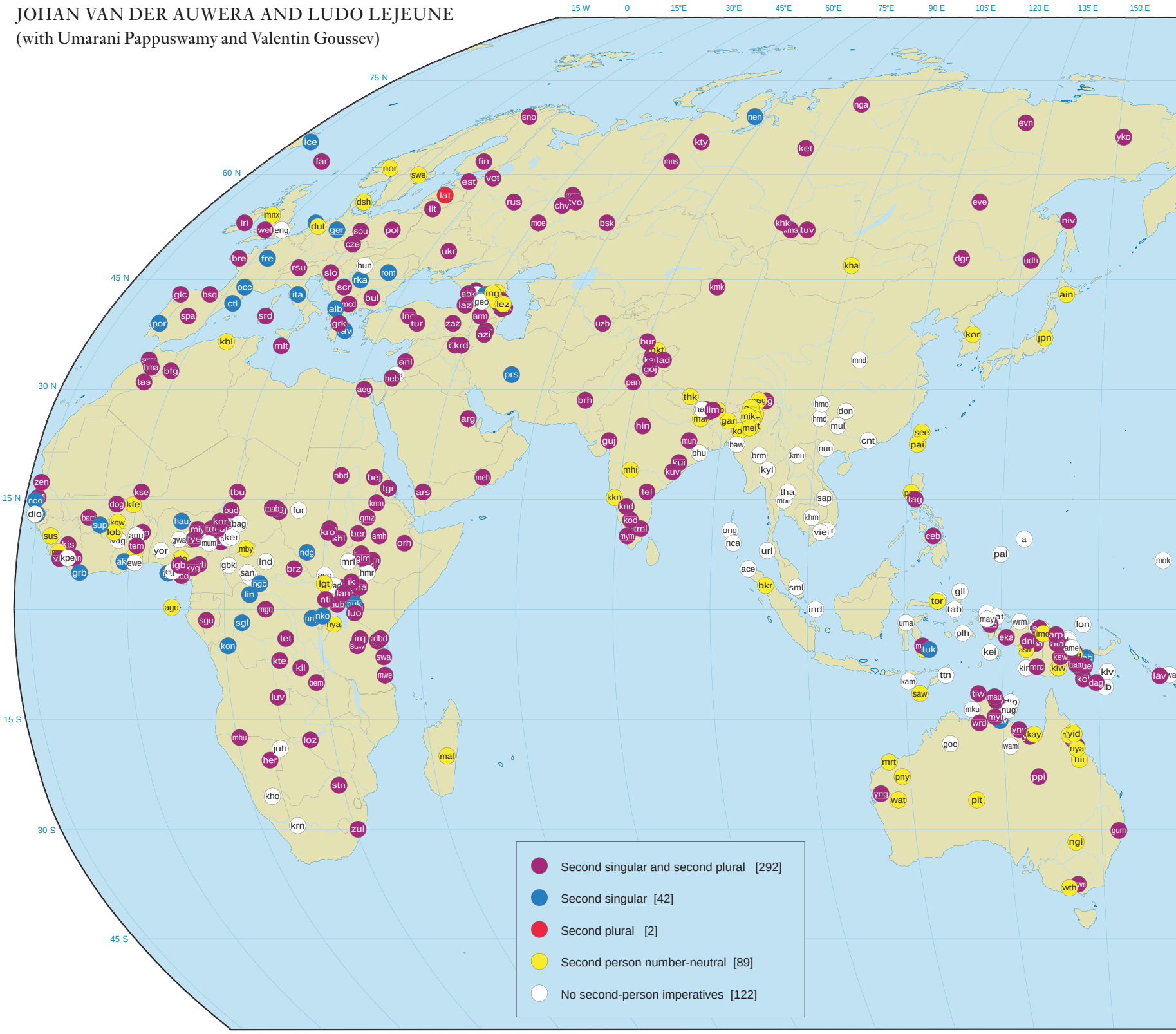
Number-specific morphological dedication comes in two subtypes. Either the second plural is morphologically dedicated or the second singular. The former is found only in Latvian and in Apurinã (Arawakan; Brazil). The latter is more common, but is on the whole concentrated in Europe, less so in West and Central Africa. Morphological dedication that is not number-specific is found in small numbers across the globe. It seems typical for the Tibeto-Burman languages of northeastern India. In Europe, it is typical for Dutch and mainland Scandinavian. The latter languages have only recently developed from the type that is most typical for Europe, viz. morphological dedication for singulars only. In some cases the imperative's insensitivity to number is in harmony with the fact that the verbs in the language are generally insensitive to number, e.g. in the Scandinavian languages or in Canela-Krahô (Macro-Ge; Brazil). But sometimes the imperative codes number to a greater extent than nonimperatives, e.g. in Nivkh (isolate; Sakhalin Island, Russia).

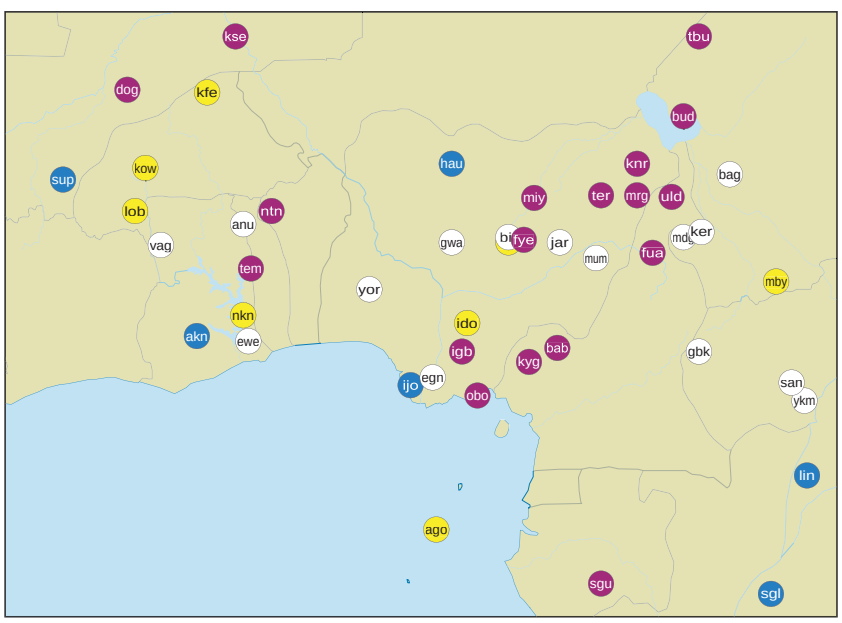
3 Theoretical issues

The morphology of the imperative has not attracted that much attention from grammarians, perhaps in part because there is relatively little morphology to start with. The most important study is Xrakovskij's *Typology of Imperative Constructions* (2001)—earlier also in Russian (1992).

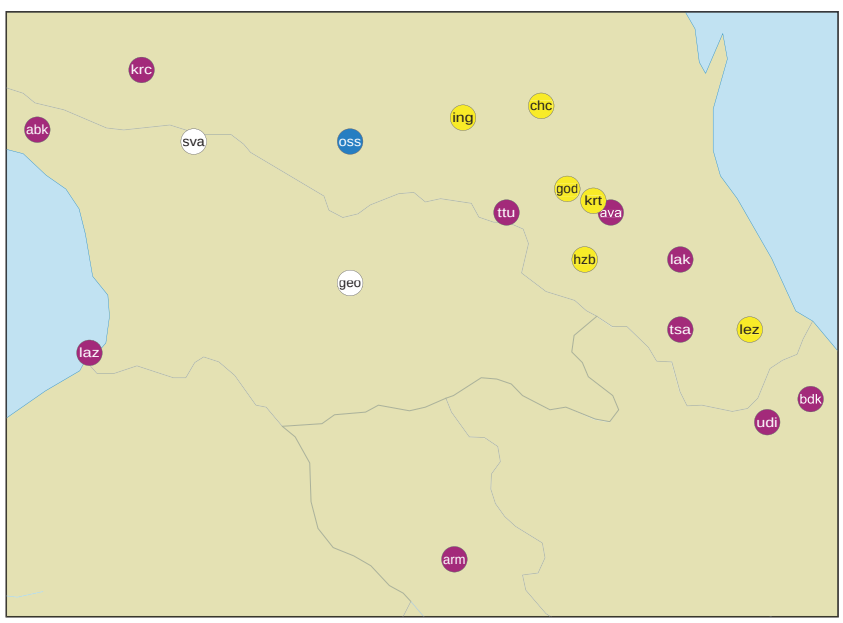
70 The Morphological Imperative

JOHAN VAN DER AUWERA AND LUDO LEJEUNE
(with Umarani Pappuswamy and Valentin Goussev)

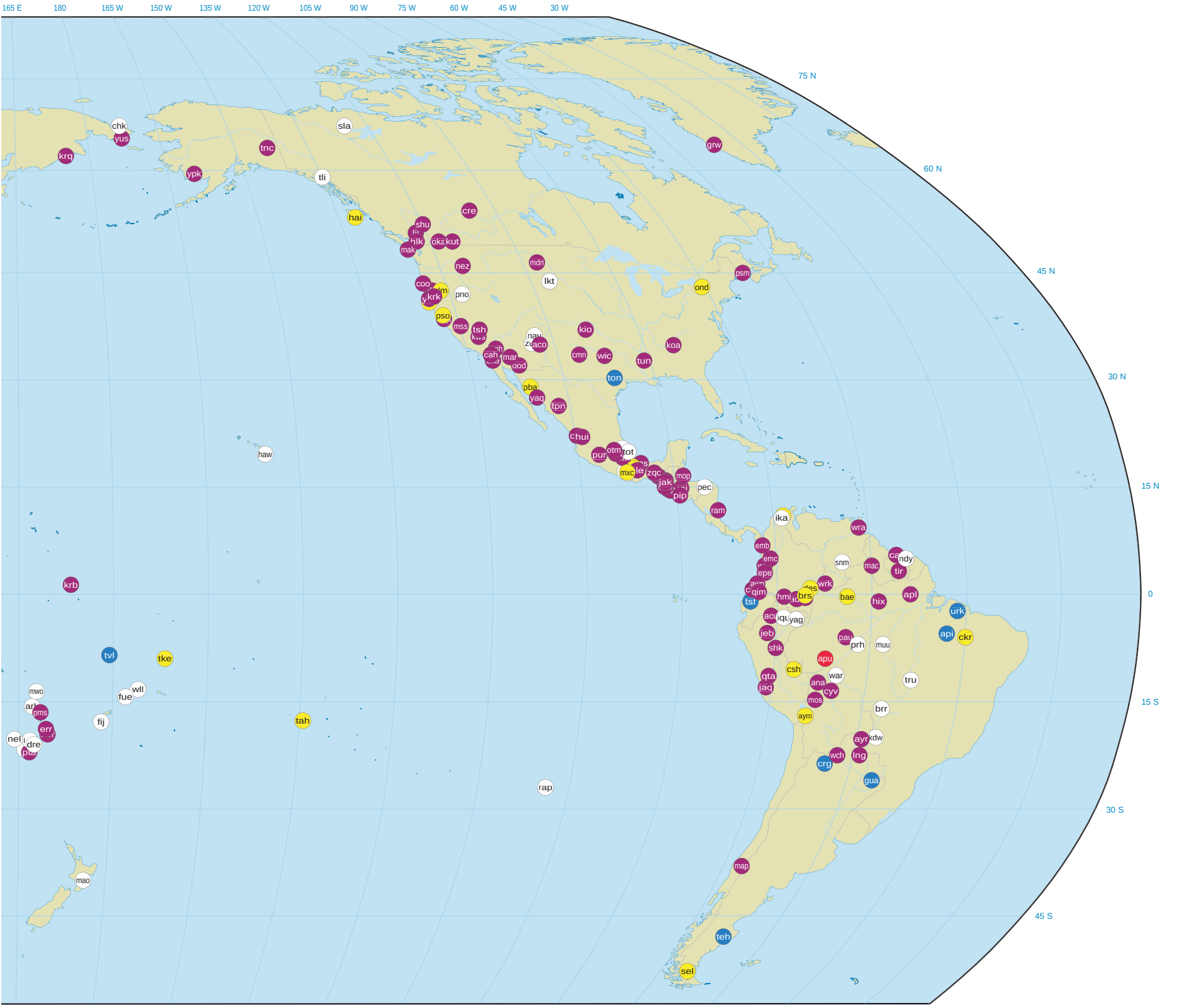


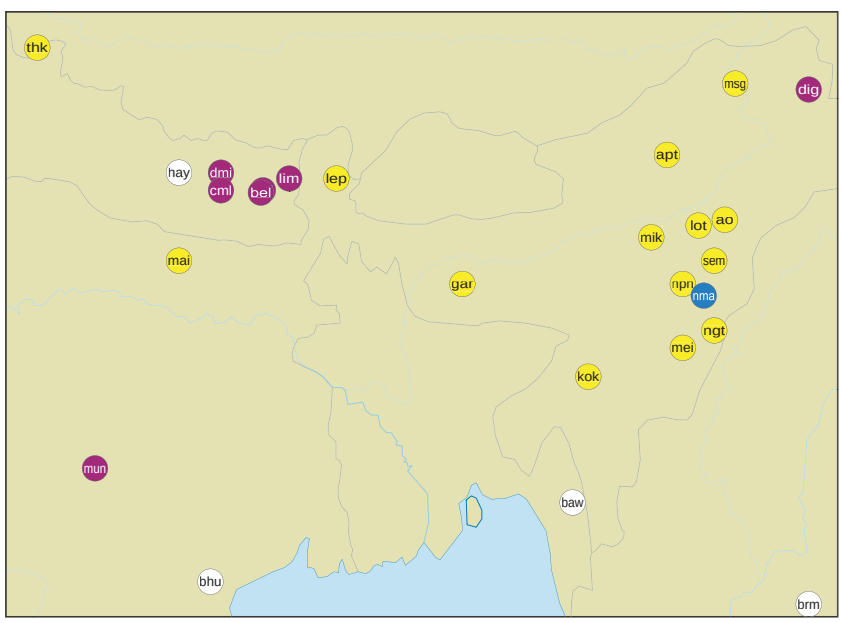


Nigeria and surrounding areas

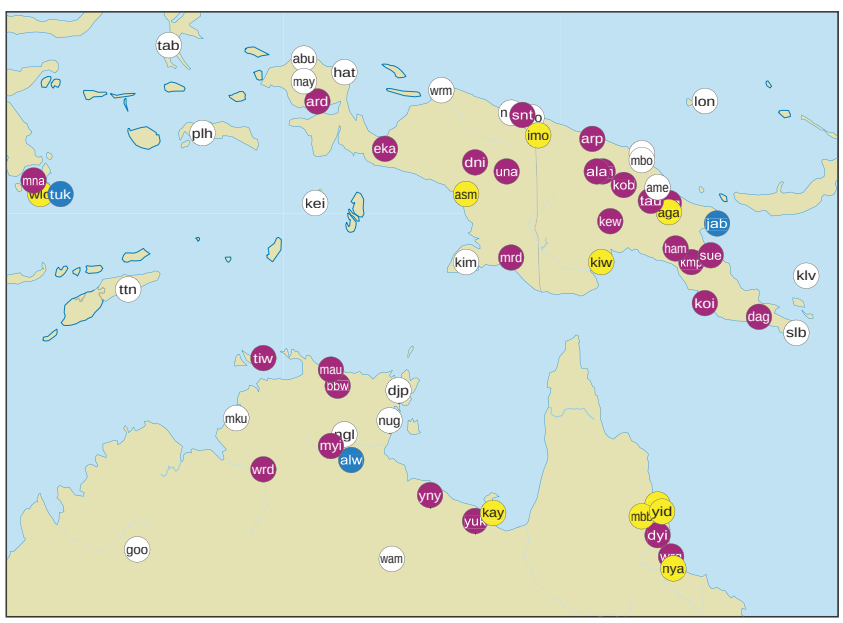


Caucasus





Nepal, eastern India, and surrounding areas



New Guinea and northern Australia

71 The Prohibitive

JOHAN VAN DER AUWERA AND LUDO LEJEUNE
(with Valentin Goussev)

1 Definition of values

This map shows the grammaticalization of a prohibition addressed to a single addressee. The category will be called the "second singular prohibitive". Four values will be distinguished.

1.	The prohibitive uses the verbal construction of the second singular imperative and a sentential negative strategy found in (indicative) declaratives.	113
2.	The prohibitive uses the verbal construction of the second singular imperative and a sentential negative strategy not found in (indicative) declaratives.	183
3.	The prohibitive uses a verbal construction other than the second singular imperative and a sentential negative strategy found in (indicative) declaratives.	55
4.	The prohibitive uses a verbal construction other than the second singular imperative and a sentential negative strategy not found in (indicative) declaratives.	144
	total	495

In the first type of language shown on the map, the second singular prohibitive is a **combination of the verbal construction used for the second singular imperative plus a sentential negative strategy found in declaratives, or in indicative declaratives, if the notion of indicative is relevant.** This type can be illustrated with Turkish.

(1) Turkish (Ceyhan Temürcü, p. c.)

a. *Okul-a* ***git!***
school-DAT go.IMP.2SG
'Go to school!'

b. *Okul-a* ***git-me!***
school-DAT go.IMP.2SG-NEG
'Don't go to school!'

c. *Okul-a* *gid-iyor-**sun**.*
school-DAT go-CONT-IND.PRES.2SG
'You are going to school.'

d. *Okul-a* *git-**m**-iyor-**sun**.*
school-DAT go-NEG-CONT-IND.PRES.2SG
'You are not going to school.'

The prohibitive second singular *git-me* in (1b) is composed of the imperative second singular shown in (1a) and the negative *me*, illustrated in the indicative declarative in (1d). As the translation indicates, English belongs to this type as well. What makes English special is that the negation not only involves the negative word *not*, but also the auxiliary *do*.

The second type employs a **second singular imperative construction, but the sentential negation strategy is not found in declaratives or, if the notion of indicative is relevant, in indicative declaratives**. In Vietnamese, declaratives are negated with *chẳng* or *không*, whereas prohibitives use the negations *chớ* and *dung*. With respect to the verbal forms used, however, imperatives and prohibitives do not differ, nor do they differ from declaratives.

(2) Vietnamese (Thompson 1965: 221; Victoria Rosén, p. c.)

a. ***Uông*** *ruou!*
drink alcoholic
'Drink alcohol!' *or* 'I/you/he/etc. is/are drinking alcohol.'

b. ***Chó*** ***uông*** *ruou!*
NEG drink alcoholic
'Do not drink alcohol!'

c. ***Không*** ***uông*** *ruou.*
NEG drink alcoholic
'I/you/he/etc. are not drinking alcohol.'

In the third type, the second singular prohibitive employs **a sentential negative that is used in declaratives or, if the notion of indicative is relevant, in indicative declaratives, but the verbal construction is not the one found in second singular imperatives**. In Spanish the sentential negation for both negative indicative declaratives and prohibitives is the marker *no*. However, the prohibitive verb is not an imperative, but a subjunctive.

(3) Spanish

a. *Pedro* ***canta.***
Pedro sing.IND.PRES.3SG
'Pedro sings.'

b. *Pedro* ***no*** ***canta.***
Pedro NEG sing.IND.PRES.3SG
'Pedro does not sing.'

c. ***Canta!***
sing.IMP.2SG
'Sing!'

d. ***No*** ***cantes!***
NEG sing.SBJV.PRES.2SG
'Don't sing!'

In the fourth type, **the verbal construction of the second singular prohibitive is different from that of the second singular imperative, and the negative strategy is different from the sentential negative found in declaratives or, if the notion of indicative is relevant, in indicative declaratives.** In Zulu (Bantu; South Africa) the imperative second singular ends in *-a*. In the prohibitive second singular, however, we do not find an imperative form but an infinitive, combined with a special negative marker *musa*, different from the indicative declarative sentential negative marker.

(4) Zulu (Poulos and Bosch 1997: 19; Khosi Mnyakeni, p. c.)

a. ***Shay-a*** *inja!*
hit-IMP.2SG dog
'Hit the dog!'

b. ***Mus-a*** ***uku-shay-a*** *inga!*
NEG.IMP.AUX-2SG INF-hit-INF dog
'Do not hit the dog!'

c. ***U-ya-shay-a*** *inja.*
2SG-IND.PRES-hit-PRES dog
'You hit the dog.'

d. ***A-wu-shay-i*** *inja.*
NEG.IND.PRES-2SG-hit-NEG.IND.PRES dog
'You do not hit the dog.'

Several complications occur. First, the prohibitive negation may be partially identical to an (indicative) declarative negation. In Yakoma (Adamawa-Ubangian; Central African Republic) there are four declarative negations and two prohibitive ones. The latter consist of two parts, and the second part is in each case identical to a declarative negation. Negative markers in Yakoma are given in example (5).

(5) Yakoma (Boyeldieu 1995: 131–2)

a. Declarative	b. Prohibitive
mɔ́ . . . mɔ̄	*tá . . . mɔ̄*
mɔ̄	
lá ̀	
āpɛ̀	*tá . . . āpɛ̀*

Partial similarity implies difference, however, and a language like Yakoma will therefore be taken to have a special prohibitive negation.

A second complication is that a language may have more than one second singular imperative. In this chapter, however, we restrict our attention to the most common or neutral imperative. Special future or special polite imperatives, for instance, are not included. This imperative, however, must have a corresponding prohibitive. Thus Kerek (Chukotko-Kamchatkan; eastern Siberia) has both perfective and imperfective imperatives, and the perfective imperative is more frequent than the imperfective imperative (Volodin 2001: 150). However, the prohibitive only has an imperfective paradigm (Volodin 2001: 153). Hence for Kerek, only the imperfective has been taken into consideration.

A third complication is that a language may possess more than one prohibitive strategy. In case the strategies are of the same type, there is no problem. In Yakoma, there are two prohibitive markers, but both are different from the declarative ones, so the language can be said to be of the special prohibitive negation type. Whether there is just one such negative or two is irrelevant. But the strategies may also be of different types. In Apalaí (Carib; Brazil), the prohibitive normally uses the copula, as in negative declaratives, and the declarative negative suffix *-pyra*. But there is also a pattern with a special prohibitive prefix *os-*, which combines with the verb in its immediate past form.

(6) Apalaí (Koehn and Koehn 1986: 62)

a. *Yto-**pyra*** ***eh-to-ko!***
go-NEG be-2PL-IMP
'Do not go (you all)!'

b. *Tupito epery **os-enah-no,*** *n-ase.*
field fruit NEG-eat-IND.IMM.PRET 3-say.IND.PRES
' "Don't eat the fruit of the field", he said.'

This second strategy is considered to be rare, however. In such cases, we have coded the language according to its most frequent pattern. For all the relevant languages on the map, we have made a frequency-based decision.

Note that we do not rule out the possibility that a language may have a special prohibitive negation which has an additional function in non-negative contexts. Iraqw (Southern Cushitic; Tanzania) distinguishes between the nonprohibitive negative *-ká* and the prohibitive *m-* (Mous 1992: 151, 168). This *m-*, however, is also used to make questions, and Mous (1992: 152) thinks that this is no accident, for there are both formal and semantic similarities.

2 Geographical distribution

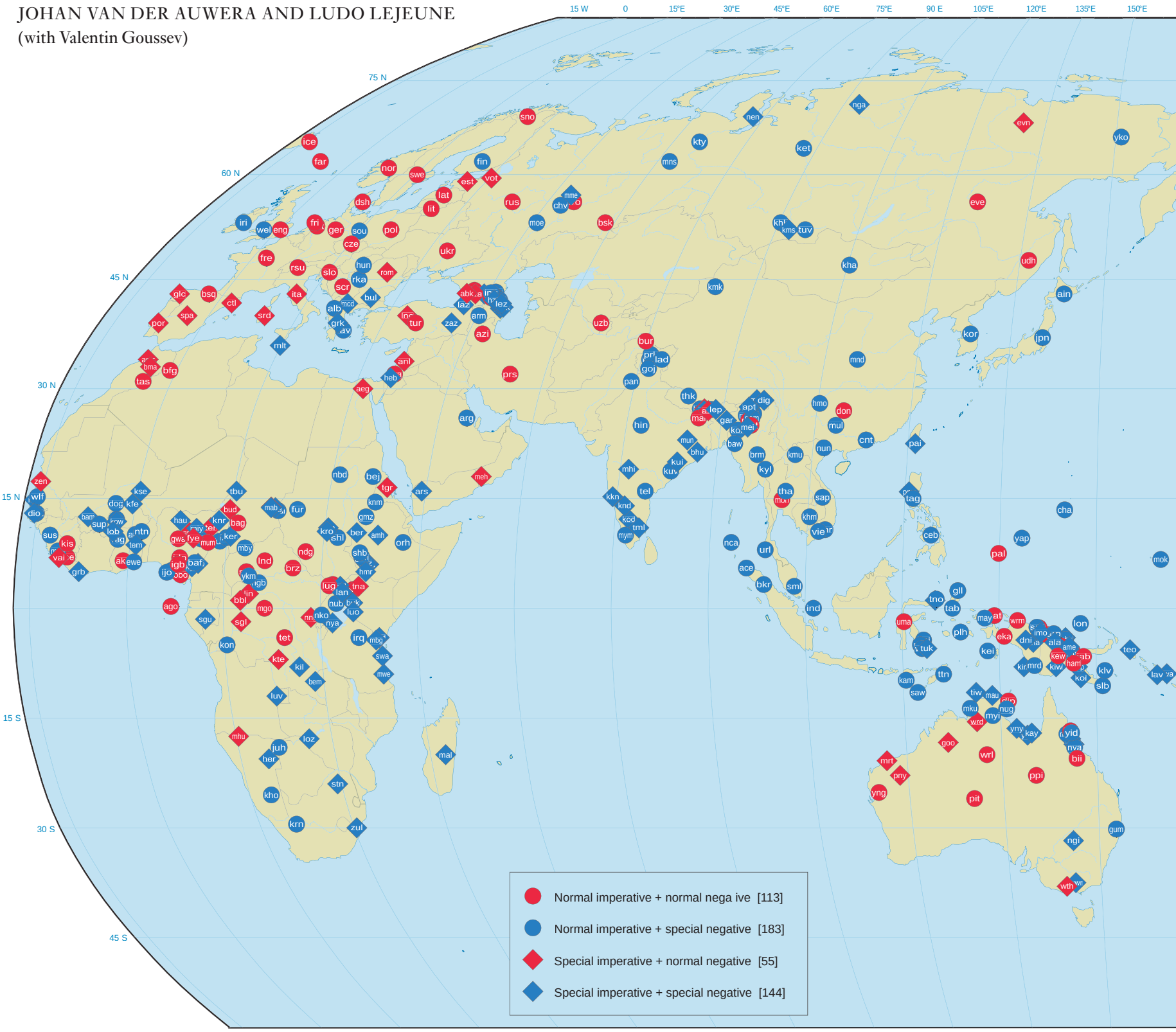
It turns out that most of the world's languages do not form second singular prohibitives with a combination of the second singular imperative and the declarative (indicative) negation. This is not to say that this pattern is rare. It is well attested in the Americas and in non-Bantu Niger-Congo; in Eurasia it is typical for Germanic, Slavic, and Turkic. However, the strategy is rare to nonexistent in South and South-East Asia and in Africa, other than in the northern non-Bantu Niger-Congo belt. Of the three remaining strategies, the one that is least common is to use the (indicative) declarative negation in combination with a verbal form that is not found in the second singular imperative. In Europe, this is typical for the Romance languages; it is sporadically found across the globe, except for South Asia, South-East Asia, and New Guinea. The strategy of combining the second singular imperative with a special prohibitive negation is widespread. It seems typical for Celtic and perhaps also Khoisan. Areally, it is most typical for South-East Asia and the Far East. Finally, it is also common to have both a special prohibitive negation and a verbal form different from the one used for the second singular imperative. This is found everywhere except in western Europe (apart from the Balkans) and in South-East Asia and the Far East. There seems to be a cluster in the southern and eastern part of India. One can also generalize over languages with a verbal construction not found in positive imperatives, whether or not they also employ a special prohibitive negation. This is typical for the Bantu languages, as has been remarked by Kamba Muzenga (1981: 256–65). A generalization about languages that have a special prohibitive negation, whether or not the verbal construction differs from the one used in second singular imperatives, holds for India. Given the relatively small number of languages investigated, it is difficult and usually impossible to make statements about smaller areas or language families. Just one illustration: with only two Arawakan languages on the map, Bare and Warekena, which both employ a special prohibitive negative strategy, we cannot improve upon Aikhenvald's (1999: 96) statement that the Arawak family as a whole tends to have a special prohibitive negative strategy.

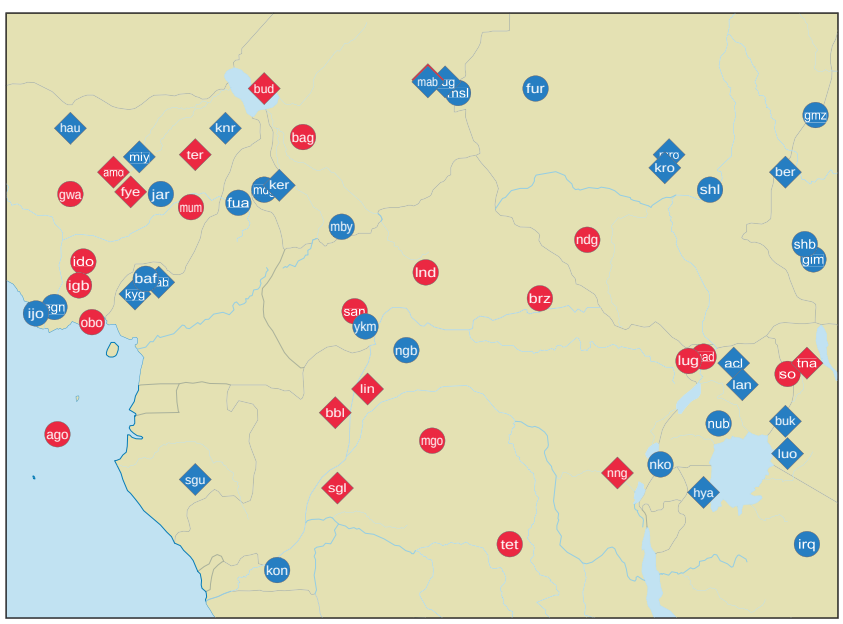
3 Theoretical issues

Though negation is a major theme in both typology and formal grammar, prohibitive negation has not figured very prominently. For typology, one should mention Kahrel and van den Berg (1994), Kahrel (1996), Hovdhaugen and Mosel (1999), and Xrakovskij (2001). Both Kahrel and van den Berg (1994) and Xrakovskij (2001) contain contributions on languages all over the world, but the sample is limited. Kahrel (1996) is a study based on forty languages and it offers some worldwide areal generalizations. Hovdhaugen and Mosel (1999) focuses on one specific language family, viz. Oceanic. Generative grammarians of the Romance languages have tried to account for why prohibitives often do not use the same verb forms as imperatives (Zanuttini 1997). Explanations are offered in terms of type and placement of negation and in terms of the morphological dedication of the imperative.

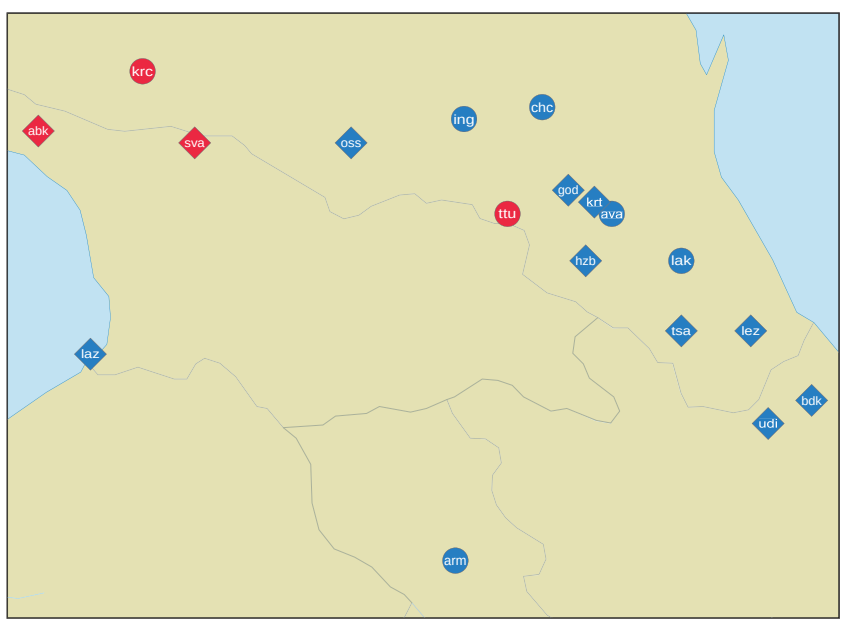
71 The Prohibitive

JOHAN VAN DER AUWERA AND LUDO LEJEUNE
(with Valentin Goussev)

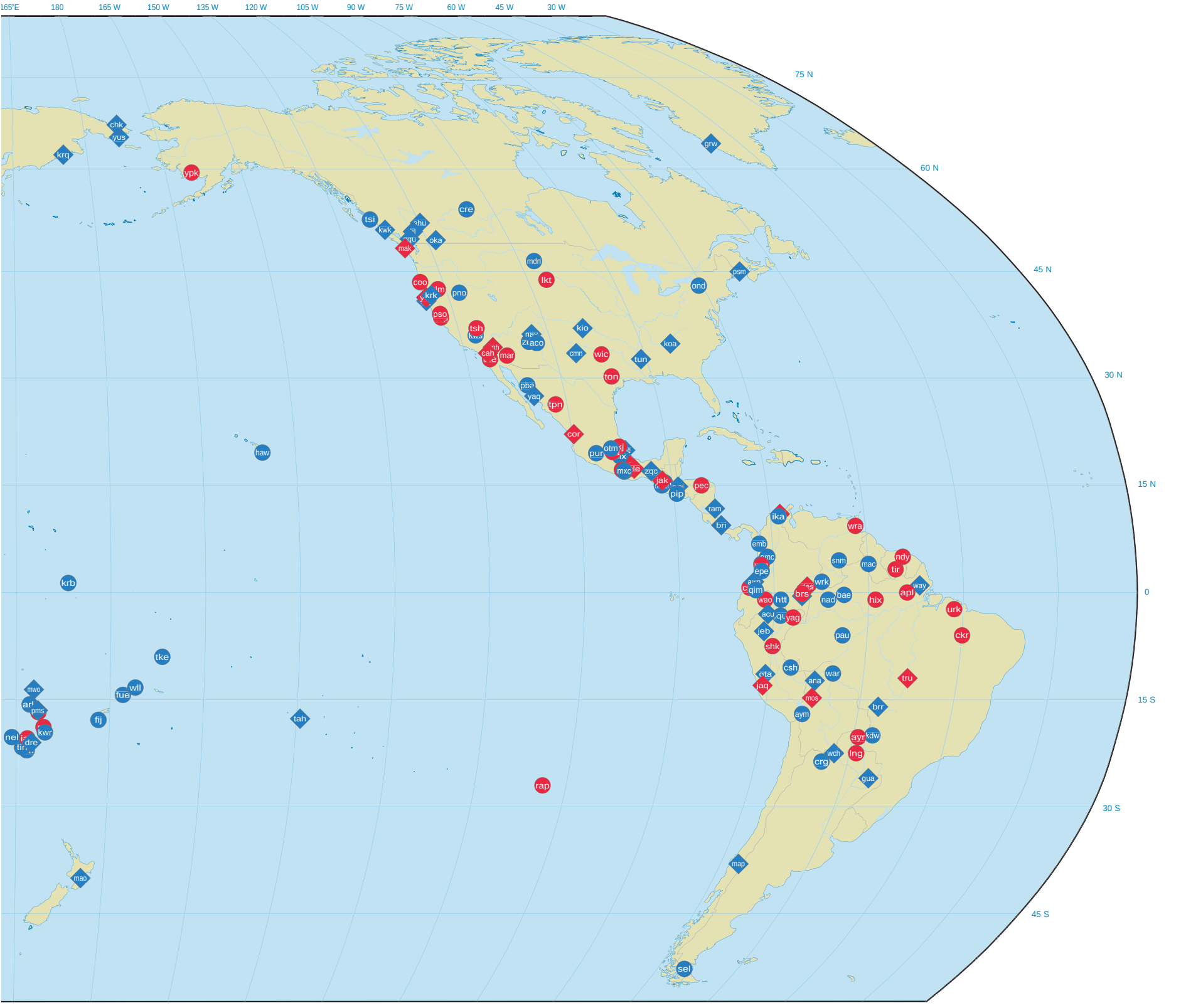




central Africa



Caucasus



New Guinea and northern Australia

Mesoamerica

72 Imperative-Hortative Systems

JOHAN VAN DER AUWERA, NINA DOBRUSHINA, AND VALENTIN GOUSSEV

1 Definition of values

This chapter documents the homogeneity of imperative-hortative paradigms. First, the notions of imperative and hortative will be defined. Then we turn to the issue of homogeneity.

1.1 Imperative and hortative. **Imperatives** and **hortatives** both have to do with the expression of a wish of the speaker about a future state of affairs. In this respect they are like **optatives** (see Chapter 73), but in contrast to optatives, they convey an appeal to the addressee(s) to help make the future state of affairs true. In case the person in control of the desired state of affairs is the addressee or addressees, we speak of an *imperative*. In any other case, we speak of a *hortative*. Consider these examples from English:

(1) a. *May he live a hundred years!* (Optative)
b. *Sing!* (Imperative)
c. *Let's sing!* (Hortative)
d. *Let him sing!* (Hortative)

(1a) illustrates an optative. It expresses a wish of the speaker, but there is no appeal to the addressee to make it true. (1b–d) also express a wish of the speaker, and in each case, there is an appeal to the addressee to help make it true. With the desired future state of affairs specified as that of someone's singing, we note that the person who is supposed to sing is/are the addressee(s) in (1b). In (1c), however, the intended singer is the addressee(s) together with the speaker. And in (1d), the intended singer is a third person.

Note that negative structures such as (2), which can be called *prohibitives* (see Chapter 71), fall outside the scope of this chapter.

(2) *Don't sing!*

The constructions studied have to be grammaticalized, but they need not be morphologically or syntactically uniquely imperative-hortative. Thus the English form *sing* is not morphologically dedicated; but the construction as a whole, without subject pronoun and with the verb in first position, may be said to be **syntactically** dedicated. Waunana (Choco; Colombia and Panama) illustrates a **morphologically** dedicated imperative second singular (see Chapter 70).

(3) Waunana (Sanchez and Castro 1977: 73)
Cö-ba!
eat-IMP.2SG
'Eat!'

An example of an imperative that is neither morphologically nor syntactically dedicated can be found in Nunggubuyu (Australian, Gunwinyguan; Northern Territory, Australia), which uses the future for the expression of imperatives.

(4) Nunggubuyu (Heath 1984: 343–4)
Ba-bura!
2SG-sit.FUT.CONT
'Sit!' *or* 'You will sit!'

One could argue that Nunggubuyu does not have an imperative, or that the form called "future" is vague between a strictly future meaning and an imperative one. In the case of Nunggubuyu, in which no other form regularly conveys the imperative meaning, the morphologically future forms will be counted as imperatives. This makes Nunggubuyu different from English. English *you will sit* will not be included as an imperative, nor will the modal construction in (5a) or the construction with a main verb *want* in (5b).

(5) a. *You should sing!*
b. *I want you to sing.*

Some cases are difficult to judge, even for a language as well studied as English. Third person *let*, for instance, can be analysed as a main verb, much like *want* in (5b). Still, Quirk et al. (1985: 148) grant third person *let* the status of a pragmatic particle, although they note that the grammaticalization of the first person plural form *let's* has progressed further, in particular with respect to its univerbation with the pronoun.

Note that in terms of grammaticalizability, grammatical persons are not equal. Cross-linguistically, the most pervasively grammaticalized imperative-hortative is the imperative second singular, the order or request to just one addressee. The imperative-hortative that is least easily grammaticalized is the hortative first singular, the 'let me' meaning with which a speaker exhorts himself.

1.2 Formal homogeneity. The map distinguishes between four types of imperative-hortative systems. The basic parameter is what we will call the formal **homogeneity** of the system, but we will also need two interlinked auxiliary notions, viz. the **maximality** or **minimality** of the homogeneity, defined in terms of homogeneity with the imperative second singular.

As to the notion of homogeneity, two imperative-hortative forms will be called *homogeneous* if they are formed using the same kind of morphological or syntactic means. The following parameters are relevant: (i) is the construction dedicated to the imperative-hortative? (ii) insofar as the strategy is morphological, is it of the same type in terms of distinctions such as base modification versus affixation or, within affixation, is the relevant affix ordered in the same position relative to the base? and (iii) insofar as the strategy is syntactic, is it of the same type in terms of the presence of e.g. imperative-hortative particles or pronouns? Note that in determinations of morphological homogeneity, we allow zero morphemes. In some cases it can be difficult to decide whether a marker is bound (morphological) or free (syntactic)—see Creissels (2000: 235, 238), who claims that for West African languages what are traditionally called pronouns should be analysed as prefixes.

As to maximality and minimality, if a language has a system with an imperative second-person singular that is not formally homogeneous with any of the other forms, then the language will be said to have a "minimal system". If, on the other hand, the second singular imperative is formally homogeneous with the other second persons, with the third persons, and with at least an inclusive first-person plural, then the language will be said to have a "maximal system". In terms of maximality and minimality the map distinguishes between four systems.

1.	The language has a maximal system, but not a minimal one.	133
2.	The language has a minimal system, but not a maximal one.	20
3.	The language has both a maximal and a minimal system.	21
4.	The language has neither a maximal nor a minimal system.	201
	total	375

The first type shown on the map includes **languages with a maximal system and no minimal one**. An example is Hungarian. A so-called "indefinite" paradigm for the verb *vár* 'wait' is shown in (6).

(6) Hungarian (Kenesei et al. 1998: 311)

	SINGULAR	PLURAL
1		*várjunk*
2	*várj(ál)*	*várjatok*
3	*várjon*	*várjanak*

Here there is complete homogeneity: all of the morphology is suffixal and none is fully dedicated, for the forms also have a subordinate subjunctive use. Note that the slot for the first singular remains empty. In fact, there is a first singular, and it is used with a hortative meaning, but then it is obligatorily accompanied by *hadd*, a particle meaning 'let'. This particle is optional for the other persons. So one could say that Hungarian even has two maximal imperative-hortative systems: one with *hadd* for all six persons, and one without *hadd* for all but the first singular.

The second type is that of **languages that have a minimal system, but no maximal one**. Consider the system of the intransitive imperative-hortatives of Soninke (Western Mande; Mali and, Senegal). The imperative second singular consists of the bare stem with no personal pronoun. Since this strategy is not found for any other person, it qualifies as a minimal system. For the other persons there is no one strategy; hence there is no maximal system. The second-person plural combines the stem with a special form of the second-person pronoun. The hortatives require the normal pronoun as well as the marker *n* or the markers *nà* and *n*.

(7) Soninke (Diagana 1995: 239)

a.	IMP.2SG:	no pronoun + stem
b.	IMP.2PL:	special pronoun + stem
c.	HORT.1PL:	normal pronoun + *n* + stem
	HORT.3SG/3PL:	normal pronoun + *n* + stem
d.	HORT.3SG/3PL:	normal pronoun + *nà n* + stem

The third type of language shown on the map has **both a minimal and a maximal system**. In Lingala (Bantu; Democratic Republic of Congo) the imperative second singular consists of the root of the verb, followed by a high tone suffix *-á*.

(8) Lingala (Meeuwis 1998: 28)
Sál-á!
work-EPV.IMP.SG
'Work!'

This constitutes a minimal system. In order to express an order or a request to more than one addressee, Lingala resorts to what Meeuwis (1998: 29) calls a subjunctive strategy; it is formally different (person-number prefix, no high tone on the suffix) and it has a wider semantic range. This subjunctive codes not only orders and requests addressed to the second plural but to all persons, including the second singular (Meeuwis 1998: 29, and p. c.). Thus Lingala also has a maximal imperative-hortative system.

(9) Lingala (Meeuwis 1998: 28, and p. c.)

	SINGULAR	PLURAL
1	*ná-sál-a*	*á-sál-a*
2	*ó-sál-a*	*bó-sál-a*
3	*tó-sál-a*	*bá-sál-a*

The fourth type of language has **neither a maximal nor a minimal system**. Meadow Mari (Uralic; Russia) has suffixal second-person imperatives. Third-person hortatives have the same structure, and Sebeok and Ingemann (1961: 21–2) list only one paradigm, comprising both the second- and third-person forms. This system is neither minimal nor maximal.

(10) Meadow Mari (Sebeok and Ingemann 1961: 21–2)

	SINGULAR	PLURAL
2	*-Ø/-t*	*-za/-sa*
3	*-še/-že/-šo/-žo/-šö/-žö*	*-(ə)št*

2 Geographical distribution

The map shows that most languages have neither a minimal nor a maximal strategy. Put positively, they have a strategy common to the second singular and plural, optionally including also the first plural or the third person, but not both the first plural and the third person. This strategy is found across the world. The only area in which it does not seem particularly common is eastern Eurasia.

The maximal strategy is also well represented. The only clear area where it is absent is central Europe. The strategy does occur in Irish, but then it skips the central area, to show up again in the Balkans—as a Balkanism (Ammann and van der Auwera forthcoming)—and in Finno-Ugric, in both the north with e.g. Estonian, and the south with Hungarian; finally, we land in eastern Eurasia, an area for which it seems most typical.

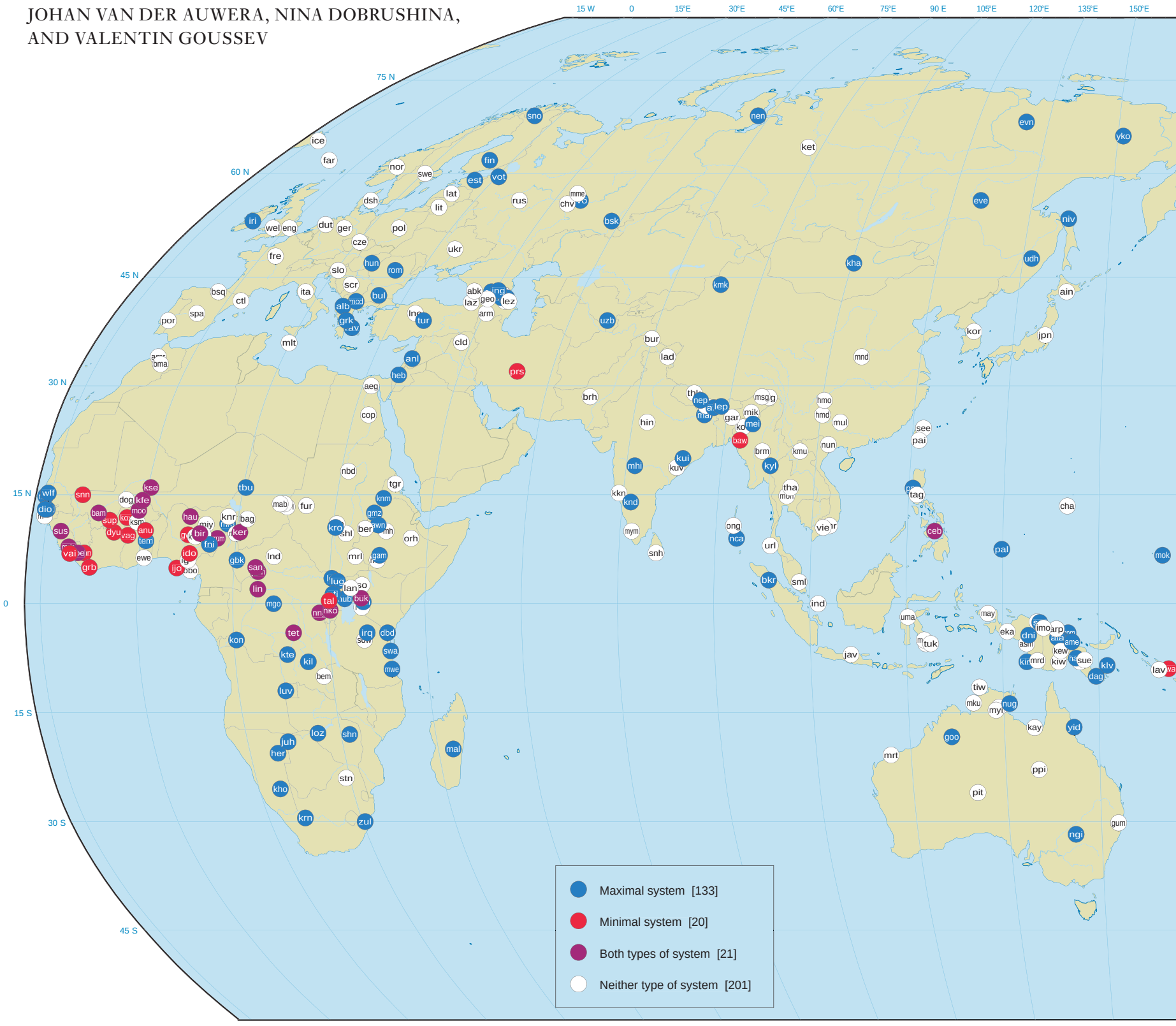
The minimal strategy is most typical for an area extending from West Africa to central Congo. There it occurs in different families of Niger-Congo: e.g. Supyire (Gur; Mali), Idoma (Benue-Congo; Nigeria), Vai (Mande; Sierra Leone), and Talinga (Bantu; Uganda). This is also the area that is typical for having both a minimal and a maximal system, this time drawing from even more widely divergent genetic groups, with, for example, Koyraboro Senni (Songhay; Mali) and Hausa (Chadic; Nigeria and Niger), and within Niger-Congo, for example, Koromfe (Gur; Burkina Faso), Grebo (Kru; Liberia), Birom (Platoid; Nigeria), Bambara (Mande; Mali), and Lingala (Bantu; Democratic Republic of Congo).

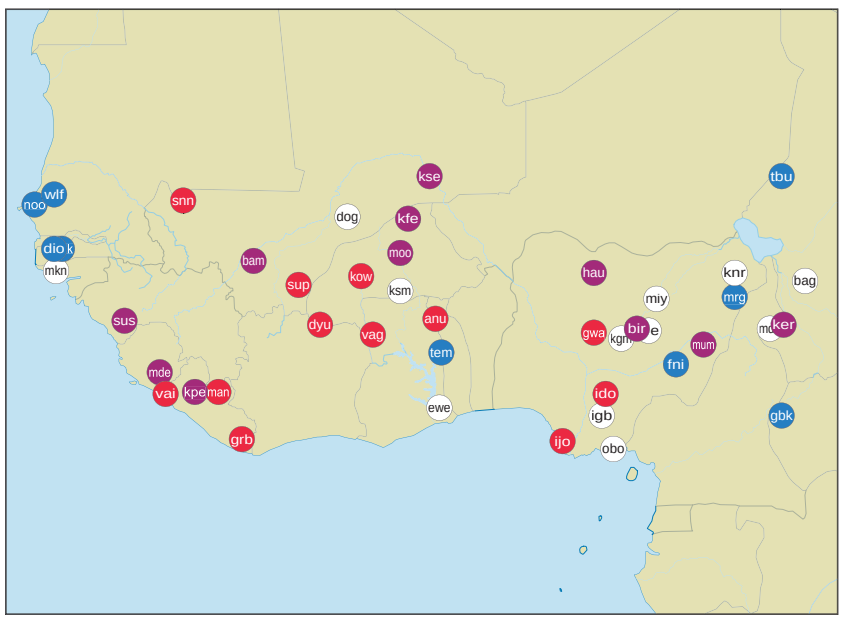
3 Theoretical issues

There are two main problems. One is how to distinguish between imperatives, hortatives, optatives, jussives, subjunctives, irrealis, and still other categories. Recent work includes Xrakovskij (2001) and Ammann and van der Auwera (2004). The other problem concerns the demarcation of imperative-hortatives from deontic modality. Some ideas can be found in Palmer (1986) and in Hengeveld (forthcoming).

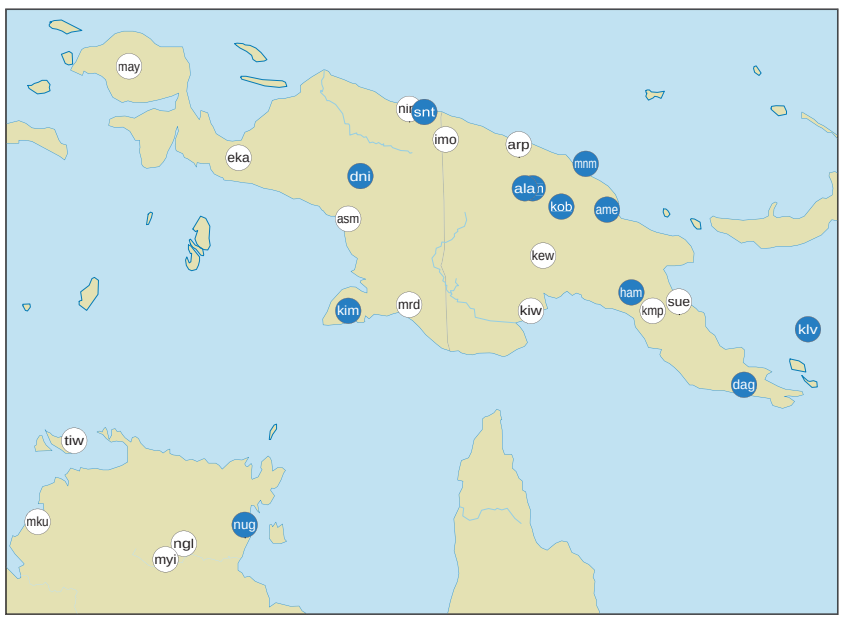
72 Imperative–Hortative Systems

JOHAN VAN DER AUWERA, NINA DOBRUSHINA,
AND VALENTIN GOUSSEV

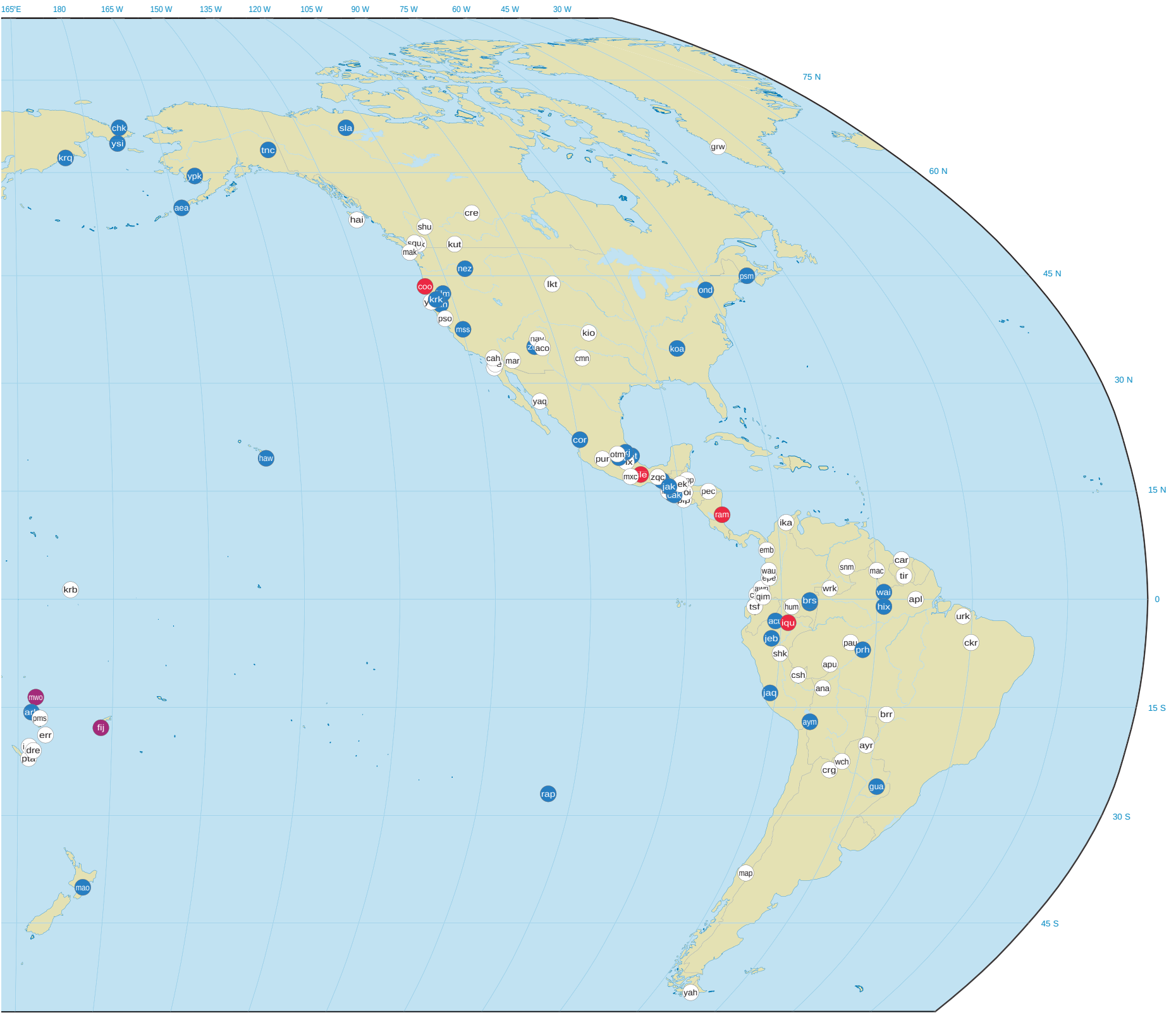


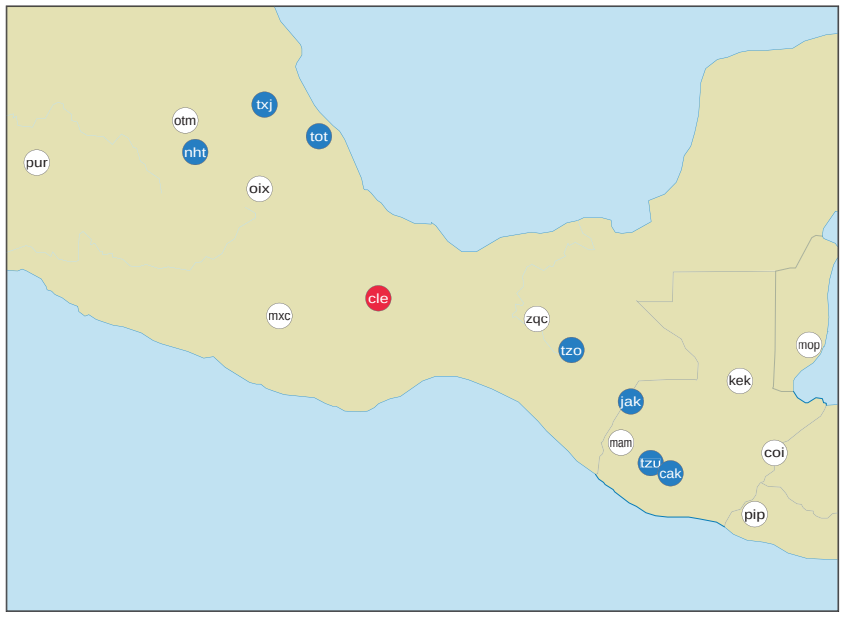


West Africa

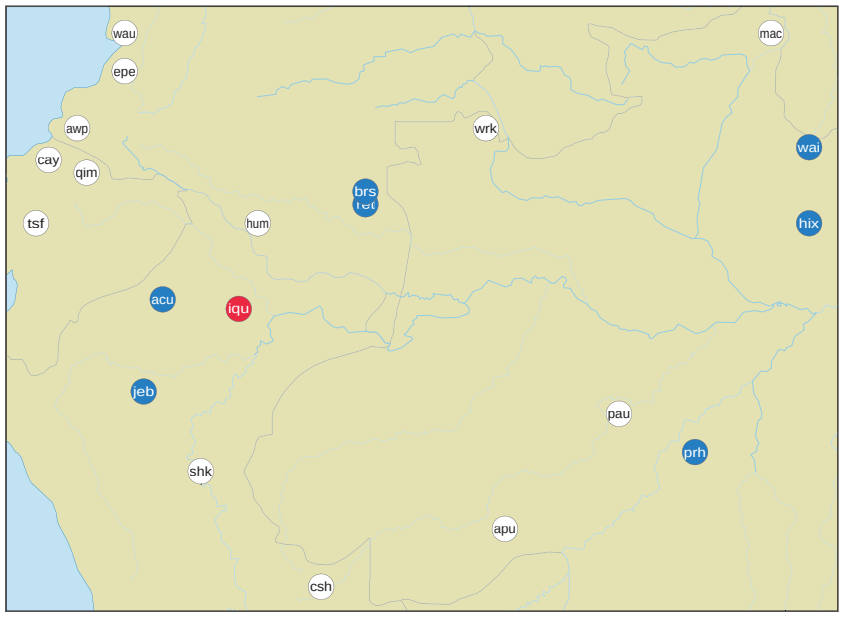


New Guinea and northern Australia





Mesoamerica



northwestern South America

73 The Optative

NINA DOBRUSHINA, JOHAN VAN DER AUWERA,
AND VALENTIN GOUSSEV

1 Introduction

In this chapter, the term *optative* will refer to an inflected verb form dedicated to the expression of **the wish of the speaker**. An example comes from Bagvalal (Nakh-Daghestanian; eastern Caucasus):

(1) Bagvalal (Dobrušina 2001: 328–9)
Den *hanč'u-b* *b-as-inā-X* *b-is-a-nā,*
1SG.ERG lying-N N-tell-IMPF-CONV N-find-POT-COND
dē *w-ič'a-be-la!*
1SG M-die-IMP-OPT
'Would that I died if I'm lying!'

This usage of the label *optative* is rather uncontroversial (Bybee et al. 1994: 179, Chung and Timberlake 1985: 247, Sadock and Zwicky 1985: 164), but other terms are also found in descriptive grammars. The speaker's wish can be expressed by different grammatical means across languages, for instance, by the use of modal verbs, such as English *may* in (2), or by a Russian elliptical conditional in (3):

(2) *May God always help you!*

(3) Russian (personal knowledge)
Jesli *by* *ona* *vernu-l-a-s'!*
if SUBJ she return-PST-F-REFL
'If only she came back!'

This chapter restricts attention to languages with a dedicated inflectional optative form as in (1), and disregards noninflectional means for expressing the speaker's wish such as (2) and (3).

2 Definition of values

The two values shown on the map are given in the feature-value box:

1.	Inflectional optative present	48
2.	Inflectional optative absent	271
	total	319

Though these values seem straightforward, in practice there are various problems. On the one hand, constructions are sometimes called "optative" which we would call something else. The construction called "optative" in Slave by Rice (1989: 1111), for instance, has a very wide range of uses, and the expression of a wish does not seem to be particularly central. Whatever this is, it is not our optative. Very striking in this respect is also the tradition in Turkology to apply the term "optative" to what we would treat as a first-person imperative or hortative (Tenišev et al. 1997: 212, 364, 402). A tendency to refer to certain imperative categories with the label "optative" is found outside Turkic grammars as well. Thus, in Fortescue's (1984: 24) description of West Greenlandic, the set of what we take to be first-person singular, first-person plural exclusive, and third-person imperatives or hortatives is distinguished from other imperative forms on formal grounds and called "optative". On the other hand, the semantic category which fits our notion of optative is often listed under another label, such as "desiderative" in Aymara (Hardman de Bautista 1974: 227), Guaraní (Gregores and Suárez 1967: 132), and Maricopa (Gordon 1986: 110), and "benedictive" in Ladakhi (Koshal 1979: 226).

It is important to note that the first value is assigned to languages that have the optative as an *inflectional* category. This means that it is not sufficient for the language to exhibit a dedicated optative construction. Consider Mao Naga (Tibeto-Burman; India). Its "blessing and cursing" optatives involve combining the particle *peno* with an imperative verb form.

(4) Mao Naga (Giridhar 1994: 314)
Pfo *peno* *thi* *mozhu-ṭio!*
he OPT die.IMP soon
'May he die soon!'

This *peno* construction is dedicated to the optative function, but it is not inflectional.

A dedicated optative form may occur in a special syntactic construction. For instance, the optative in Kumyk (Turkic; eastern Caucasus) is a combination of a participle in *-gaj/-ɣaj* and an auxiliary verb *edi*:

(5) Kumyk (Gadžiaxmedov 2000: 246)
Jaxšy *jangur* *jav-ɣaj* *edi.*
good rain rain-OPT COP
'Would that it rained!'

The strategy is analytic, but we would nevertheless say that Kumyk has optative morphology: the suffix *-gaj/-ɣaj* is fully dedicated to the expression of optatives—it is indeed the optative marker most typical for Turkic in general. Until recently, the Kumyk optative occurred without an auxiliary, but nowadays the pattern without auxiliary sounds archaic. Even so, the need for the auxiliary does not change the fact that there is morphology that is specifically optative in meaning.

Another restriction is that we only consider forms that are available for all persons. For instance, although Bashkir (Turkic; Russia) has a second-person optative to express a wish that something bad should happen (curse) (Juldašev 1981: 288), it is shown on the map as lacking an optative.

3 Related categories

It is important also to distinguish the optative from related categories. We will here discuss subjunctives, desideratives, and imperatives (hortatives).

3.1 ***Subjunctive.*** The notion of **subjunctive**, first of all, is itself a problematic notion. One can choose to define it as a mood that is typical for subordinate clauses, or at least some subtypes, or one can associate it with non-actuality and thereby bring it closer to notions like "irrealis" and "potentialis". Either way, there is a link with the expression of wishes. First, with Russian we have just illustrated in (3) the fact that the optative may use subordination strategies, and the very combination of a past tense and the particle *by* seen in (3) has been called "subjunctive". Of course, this Russian subjunctive is not a morphological category, and the Russian facts pose no problem for us. The link between the expression of wishes and the subjunctive category is, however, independent of whether the subjunctive is realized by a syntactic pattern as in Russian or by morphology only, as in Eastern Armenian.

(6) Eastern Armenian (Kozintseva 2001: 247)
Du *oɣǰ* *mnas,* *aɣǰik-s.*
you well remain.SUBJ.FUT.2SG daughter-POSS
'I wish you good health, daughter.'

To distinguish the morphological optative—the topic of this chapter—from the wish-expressing morphological subjunctive, such as that of Armenian (6), we stipulate that the expression of the wish must be the main function. This is sometimes clear enough; but when the category called "optative" has other uses, too, the

difference between a multifunctional optative and a multifunctional subjunctive may become very small, and grammarians may disagree. Swahili, for instance, is commonly taken to have a multifunctional subjunctive—Wilson (1970: 118–21) lists as many as seven functions—but Loogman (1965: 206–9) speaks of the "Wishing-Form". Multifunctional optatives used in non-wishing contexts are illustrated by (7) from Archi (Nakh-Daghestanian; eastern Caucasus) (a conditional use) and by (8) from Maricopa (Yuman; Arizona) (what Palmer (1986: 110) calls a "deliberative question" use).

(7) Archi (Kibrik et al. 1977: 227)
Noc' *sas* *eīit'u-īan,* *immaqº'i.*
bird catch be.able-OPT leave.IMP
'If you can't catch the bird, just leave it.'

(8) Maricopa (Gordon 1986: 110)
Kawish *'-we-lya?*
what 1-do-OPT
'What can I do?'

3.2 ***Desiderative.*** Another notion related to optative is **desiderative** (Plungian 2000: 317) (see also Chapter 124). Both categories involve the expression of a wish, but whereas the optative is used for the wish of the speaker, the desiderative is used for the wish of a participant in the state of affairs referred to in the utterance. A language with a dedicated inflectional strategy for the desiderative is Nganasan (Samoyedic, Uralic; northern Russia).

(9) Nganasan (Valentin Goussev, field notes)
D'ojbaru *i-n'antu-g-tu-m/-Ø.*
orphan be-DESID-ITER-PRES-1SG.SBJV/–2SG.SBJV
'I/you want to be an orphan.'

3.3 ***Imperative.*** The third related notion is the **imperative** (see also Chapter 72). Both the imperative and the optative refer to a wish of the speaker. With the optative, the state of affairs wished for is typically outside the sphere of influence of the speaker; see (10a) from Aghul (Nakh-Daghestanian). With an imperative, however, the speaker launches an appeal to the hearer to fulfil the wish. In Aghul, the optative simply cannot be used in this situation, see (10b).

(10) Aghul (Solmaz Merdanova, p.c.)
a. *Wun* *žehlem-di-s* *ušu-raj!*
you hell-OBL-DAT go-OPT
'Go to hell!'
b. **Wun* *mexteb-i-s* *ušu-raj!*
you school-OBL-DAT go-OPT
'Go to school!'

In other languages, the optative is allowed, but the appeal still cannot be a simple command. In Santali (Munda; India; data obtained with the help of Dmitrij Sitchinava), optatives are allowed for addressing the hearer and the effect is that of a mild suggestion.

(11) Santali (Bodding 1952: 73)
Mit' *taka-m* *em-k-iñ-a?*
one rupee give-OPT-1SG-IND
'Shouldn't you give me a rupee?'

The problem is that there may be an overlap between optatives and imperatives or hortatives in the third person. Third-person imperative (hortative) may be used as third-person optative and vice versa. Consider the uses of the third-person imperative in Mapudungun (Chile). A true imperative use is shown in (12a) and a semantically optative one is shown in (12b).

(12) Mapudungun (Smeets 1989: 234)
a. *Chewrumé* *rüngkü-pe!*
wherever jump-IMP.3
'Let him jump wherever he wants to!'
b. *Küme-le-pe* *ñi* *fochüm!*
good-STAT-IMP.3 my son
'May my son be all right!'

In Lealao Chinantec (Oto-Manguean; Oaxaca, Mexico; Rupp 1989: 95) and in Lezgian (Nakh-Daghestanian; Haspelmath 1993: 151–2), on the other hand, it is the optative third person that does double duty.

4 Further issues

There are languages with more than one morphological optative, and in that case they will have different meanings. Thus in Abkhaz (North-West Caucasian), one of the optatives, glossed as OPT1 here (Hewitt's (1979: 198) "subjunctive"), is used to curse and to bless, while the other one, the OPT2, is used to express a wish or a dream of the speaker.

(13) Abkhaz (Hewitt 1979: 198; Hewitt, p.c.)
a. *Bzia* *wy-ba-aait'!*
good you-see-OPT1
'Greetings!' (lit. 'May you see something good!')
b. *Ma* *a-dzə* *l-zhʷə-nda.*
PCL DEF-water 3.F-drink-OPT2
'I wish she'd drink the water.'

5 Geographical distribution

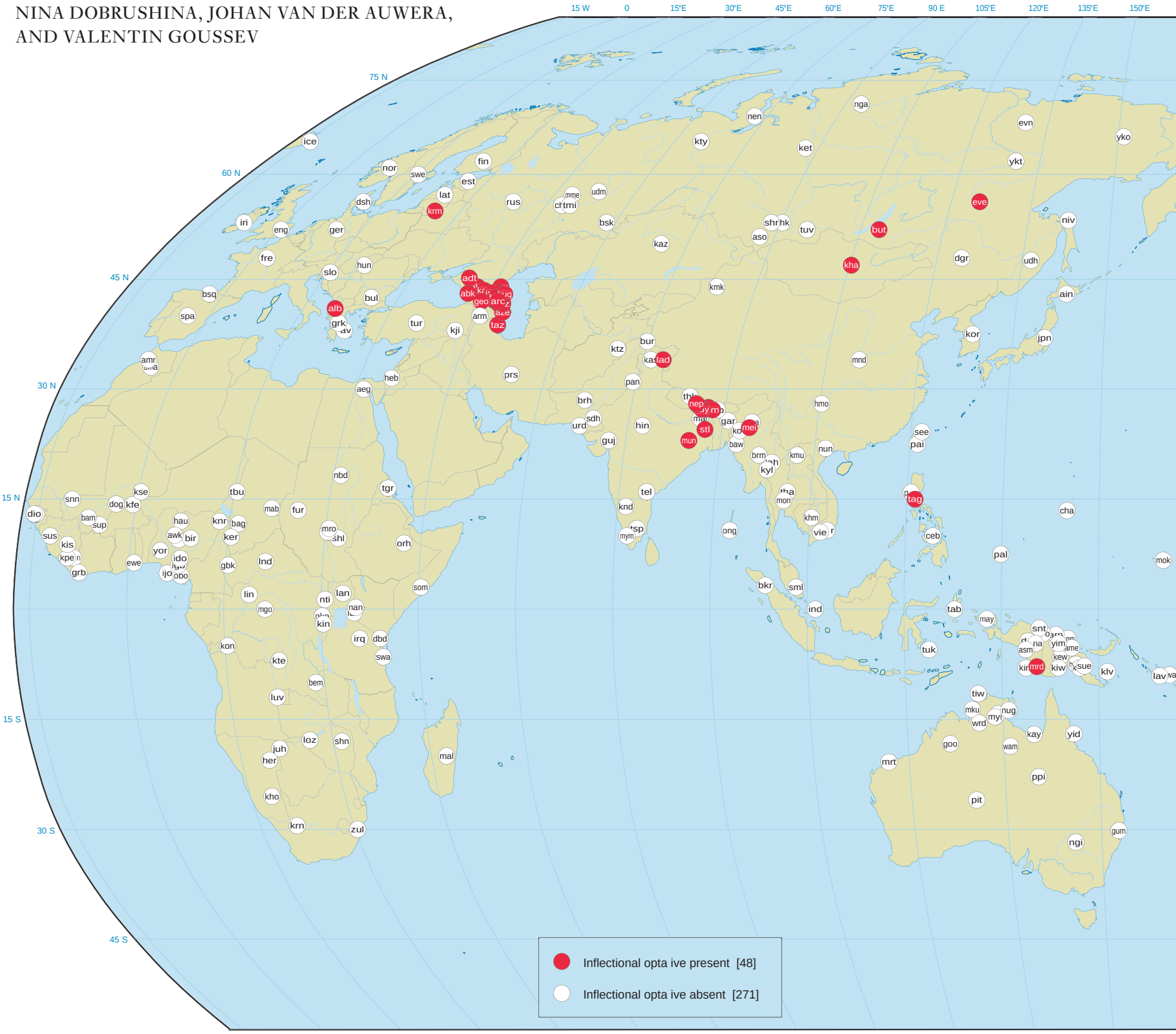
The map documents 319 languages. Worldwide, so it seems, morphological optatives are a fairly infrequent phenomenon (48 languages in our sample). Our sample does not contain any languages with inflectional optatives from Australia or from Africa; we have just two cases in Europe (Albanian and Karaim), one in New Guinea, and only a few in the Americas and Asia outside of the Caucasus and the area of northern India and Nepal.

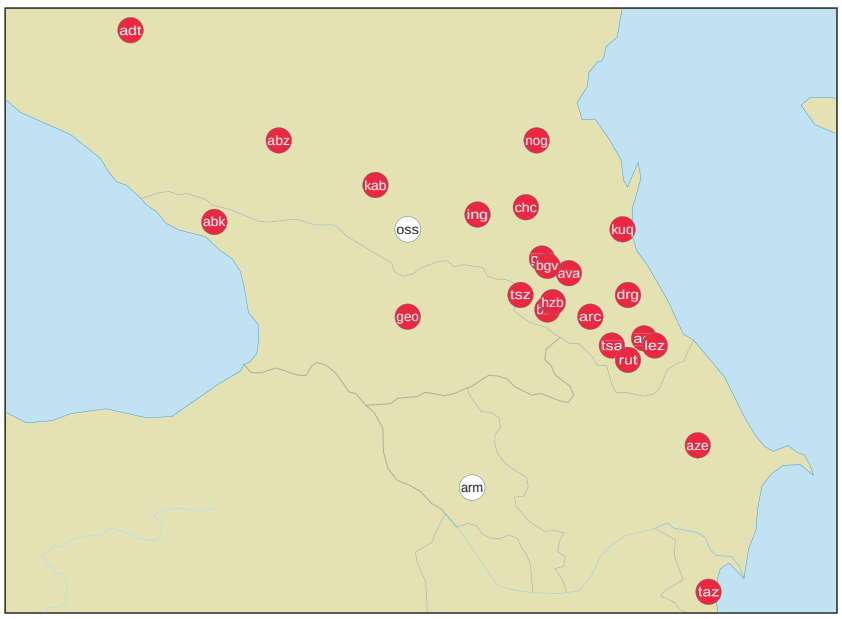
As far as the Caucasus is concerned, the optative was found in all of the North-Caucasian languages for which we were able to find reliable information. North-West Caucasian languages, furthermore, seem highly unusual in that they have two optatives; Ingush and Chechen, which are North-East Caucasian (= Nakh-Daghestanian), also have two optatives, though one of them is unproductive (Johanna Nichols, p.c.). Unfortunately, data from the South Caucasian (= Kartvelian) languages proved less accessible. Georgian, at least, does have a morphological optative. Whether the Caucasian optative is a phenomenon due to areal contact or to a common origin is not clear, in part because the genealogical affinity of the North-West and North-East languages is still under discussion. Interestingly, both Turkic languages of the Caucasus, Karachay-Balkar and Kumyk, have a full personal paradigm of a morphologically dedicated optative, which is atypical for Turkic.

A second area that seems to be characterized by the morphological optative is northern India and Nepal. The optative occurs in some of the Tibeto-Burman and Munda languages, and interestingly, the only Indo-Iranian language we found to have a morphological optative is Nepali (Korolev 1965: 80), probably as a result of the influence of neighbouring Tibeto-Burman languages.

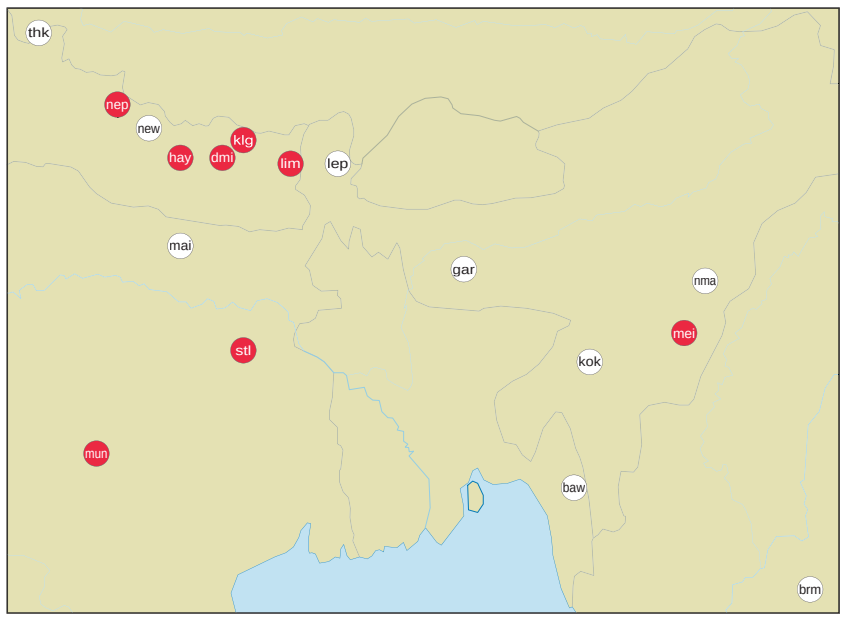
73 The Optative

NINA DOBRUSHINA, JOHAN VAN DER AUWERA,
AND VALENTIN GOUSSEV

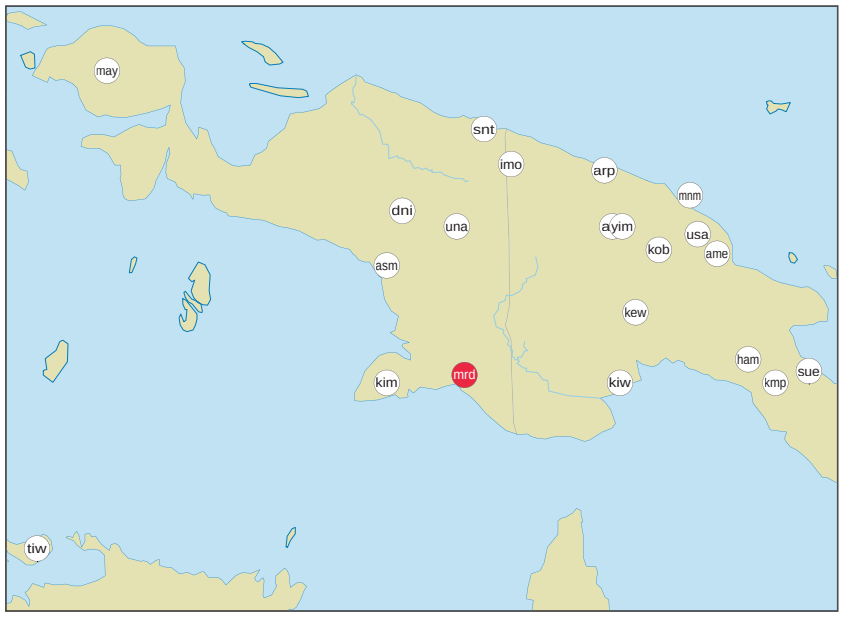




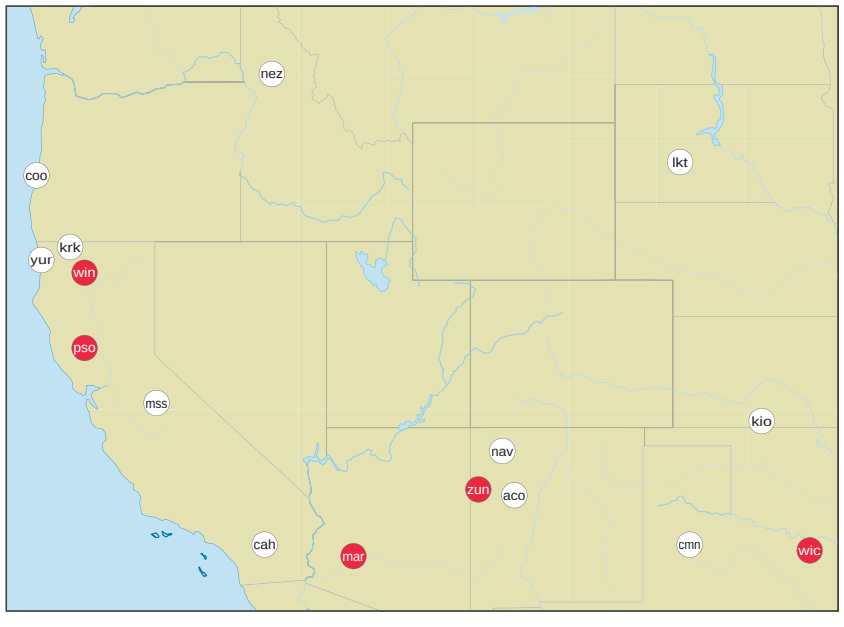
Caucasus



Nepal, eastern India and surrounding areas



New Guinea and northern Australia



western United States

74 Situational Possibility

JOHAN VAN DER AUWERA AND ANDREAS AMMANN

1 Definition of values

In this chapter, we will examine the expression of the equivalents of English *may* and *can* as in (1).

(1) a. *The children* ***can*** *swim across the lake.*
b. *You* ***may*** *leave now.*
c. *One* ***can*** *get to Staten Island using the ferry.*

In (1a) *can* expresses ability and in (1b) *may* expresses permission. Both permission and ability have to do with a notion of **possibility**. These two kinds of possibility belong to a subtype which also includes readings like *can* in (1c), where neither an individual's ability nor an authority granting permission is involved. What the three examples have in common is that the speaker merely describes a possibility that exists in a given situation. For this reason we call this subtype **situational** (also termed "deontic" or "root" in the literature). It is different from the kind of possibility illustrated in (2).

(2) *John* ***may*** *have arrived.*

In (2) the possibility concerns the speaker's evaluation of the truth value of a proposition about a situation, that is, it expresses a degree of the speaker's certainty. This kind of possibility is commonly called **epistemic** (see Chapter 75).

Possibility stands in direct contrast with **necessity**, and together they constitute the domain of **modality**. Just as there are two types of possibility, there are two types of necessity, a situational one, as in (3a), and an epistemic one, as in (3b).

(3) a. *John* ***has to*** *leave now.*
b. *John* ***must*** *have arrived.*

The map gives some indication about the kind of strategy used to express situational possibility in positive main clauses. Three types of languages are distinguished. The definitions make reference to three marking strategies: affixes on verbs, verbal constructions, and other constructions, ordered in terms of increasing cross-linguistic frequency. The first type on the map comprises languages that use the least frequent strategy whether or not they also use the more frequent ones. Languages of the second type use the medium frequency (but not the low frequency), whether or not they also use the most frequent ones. Languages of the third type use only the most frequent strategy.

1. The language can express situational possibility with affixes on verbs.	63
2. The language does not express situational possibility with affixes on verbs, but with verbal constructions.	158
3. The language does not express situational possibility with affixes on verbs or with verbal constructions, but with other kinds of markers.	13
total	234

If one compares the frequency of the types of markers and the frequency of the types of languages, the most remarkable fact is that languages of type three, which use only the most frequent type of marking, are not very frequent.

The first type includes **languages that express situational possibility with verbal affixes**. An example is Lavukaleve (Solomons East Papuan; Solomon Islands) and the relevant affix is *-nen*~*-nan*. In the context for sentence (4), the speaker has just heard that a feast is planned.

(4) Lavukaleve (Terrill 2003: 344)
Valai! *Ngai* *a-vo-****nen****.*
how I 1SG.SUBJ-come-POS
'What? Can I come?'

Verbal affixes may differ in their degree of dedication to the modal meaning. "Less dedicated" means that the marker has other meanings, or—in part depending on the point of view of the grammarian—that the marker has only one meaning but various uses, with a meaning that is vague and thus wider than just situational possibility. These other meanings or uses are often future or irrealis: if you can or must do something, you normally have not done it yet. An example of a morphological verbal marker that denotes both future and situational possibility is shown from Daga (Dagan; Papua New Guinea).

(5) Daga (Murane 1974: 58)
*War-****ait****.*
get-FUT.1PL
'We will get (it).' or 'We can get (it).'

It is not always easy to decide whether a marker is a verbal affix or a free word, a verb or a particle, for instance. For Gadaba (Dravidian; India) the element *-occuni* is considered to be an invariant modal (subjunctive) suffix by Bhaskararao (1998: 349), yet it is written as a separate word. Since we take it that it is more interesting to draw attention to a low frequency phenomenon (affixal coding) than a higher frequency phenomenon (verbal construction), Gadaba is listed as a language that can express situational possibility with an affix on the verb.

The second type includes **languages that do not express situational possibility with verbal affixes, but that employ verbal constructions.** This type is well known from Standard Average European. French *pouvoir* 'may/can' is clearly a verb, appearing as a first person singular present in (6a), as a second plural future in (6b), and as a past participle in a 'have' perfect in (6c).

(6) French
a. *Je* ***peux*** *le lui donner.*
I can.IND.PRES.1SG it him give
'I can give it to him.'
b. *Vous* ***pourrez*** *me visiter.*
you.PL can.IND.FUT.2PL me visit
'You will be able to visit me.'
c. *J'* *ai* ***pu*** *le voir.*
I have.IND.PRES.1SG can.PRET.PTCP him see
'I was able to see him.'

In Taba (Austronesian; Halmahera, Indonesia), the modal is a verb, too, but it differs from French in that it can take part in a **serial verb** construction, i.e. a concatenation of two finite verbs. In (7) the "modal evaluator" *-ahan* 'to be able' is juxtaposed to a lexical verb which bears the same inflection marker (that of the third person singular).

(7) Taba (Bowden 2001: 316)
N = *pe* *n* = ***ahan.***
3SG = do 3SG = be.able
'He can do it'.

Then there are languages in which the modal verbs have reduced verbal properties—or special ones—so that grammarians may assign them labels like "irregular verb" or "auxiliary verb" and thus distinguish them from regular and lexical verbs. Criteria for such a status, however, differ from language to language and therefore no such distinction will be made here.

Not all verbal constructions that can express situational possibility are dedicated to this function. French *pouvoir*, for instance, can also serve to convey epistemic possibility (see Chapter 76).

In the third type, the **language has neither verbal affixes nor verbal contructions to express situational possibility, but there are other markers.** These markers may be particles,

adverbials, nouns, or adjectives, as well as some more complex clausal constructions. Chukchi (Chukotko-Kamchatkan; eastern Siberia) makes use of a particle *mecənkə* for the expression of situational possibility.

(8) Chukchi (Dunn 1999: 76)
Mecənkə *mət-ra-jalɣət-ɣʔa.*
enough 1PL-FUT-move.camp.TH
'We can move camp.'

An adverbial marker is illustrated with Slovene.

(9) Slovene (Derbyshire 1993: 109)
Lahkó *me* *poklíčete,*
easily me call.IND.PFV.PRES.2PL
kàdar *kóli* *želíte.*
when ever wish.IND.IMPF.PRES.2PL
'You may call me whenever you wish.'

An adjectival use is exemplified with the adjective *ihoc* 'able' from Amele (Madang; Papua New Guinea) in (10).

(10) Amele (Roberts 1987: 265)
Ija *nu-ec* *nu* ***ihoc.***
I go-INF for able
'I am able to go.'

In Korean (11), it is a noun that is to be associated with the situational possibility meaning.

(11) Korean (Sohn 1994: 348)
Halapeci-nun *ɯuncenha-si-l*
grandpa-TOP drive-SUBJ.HON-PRES
sɯu(-ka) *iss-usi-ta.*
way(-NMLZ) exist-SUBJ.HON-DECL
'My grandpa can drive.'

Note that the possibility marker in (11) is strictly speaking not the noun on its own, but the combination of the noun and the existential copula. In this sense, one could argue for coding it as belonging to the verbal construction type. We will not do this, however: in (11) it is not so much the copula that has a modal sense, but the noun.

For an example of a complex clausal type we can also turn to Korean. (12) literally says that even if you go, it is good.

(12) Korean (Sohn 1994: 348)
Ne-nun *ka-**to*** ***coh**-ta.*
you-TOP go-even.if good-DECL
'You may go.'

Just as with modal affixes and modal verbal constructions, the markers of the third type need not be fully dedicated to the expression of situational possibility. The Chukchi particle *mecənkə*, for instance, illustrated in (8), can also mean 'enough' (Dunn 1999: 76, and p. c.).

2 Geographical distribution

Languages that cannot express situational possibility in the verbal domain, whether with separate verbal constructions or with affixes on verbs, seem rare. They are not reported for Europe and most of Eurasia, nor for Africa, South Asia, or South America. Elsewhere they are scattered.

Languages that may express situational possibility with verbal affixes are more frequent, particularly in North America and in New Guinea. It is not surprising that this group includes polysynthetic languages, among them West Greenlandic (Eskimo), Kiowa (Kiowa-Tanoan; central United States), Nunggubuyu (Gunwinyguan; Northern Territory, Australia), Yimas (Lower Sepik-Ramu; Papua New Guinea), and Ket (Yeniseian; Russia). Note, however, that the existence of high polysynthesis need not entail expressing situational possibility with verbal affixes. Thus Chukchi or Lakhota (Siouan; Nebraska and Minnesota), both having a high degree of synthesis, do not seem to choose verbal affixes for expressing situational possibility.

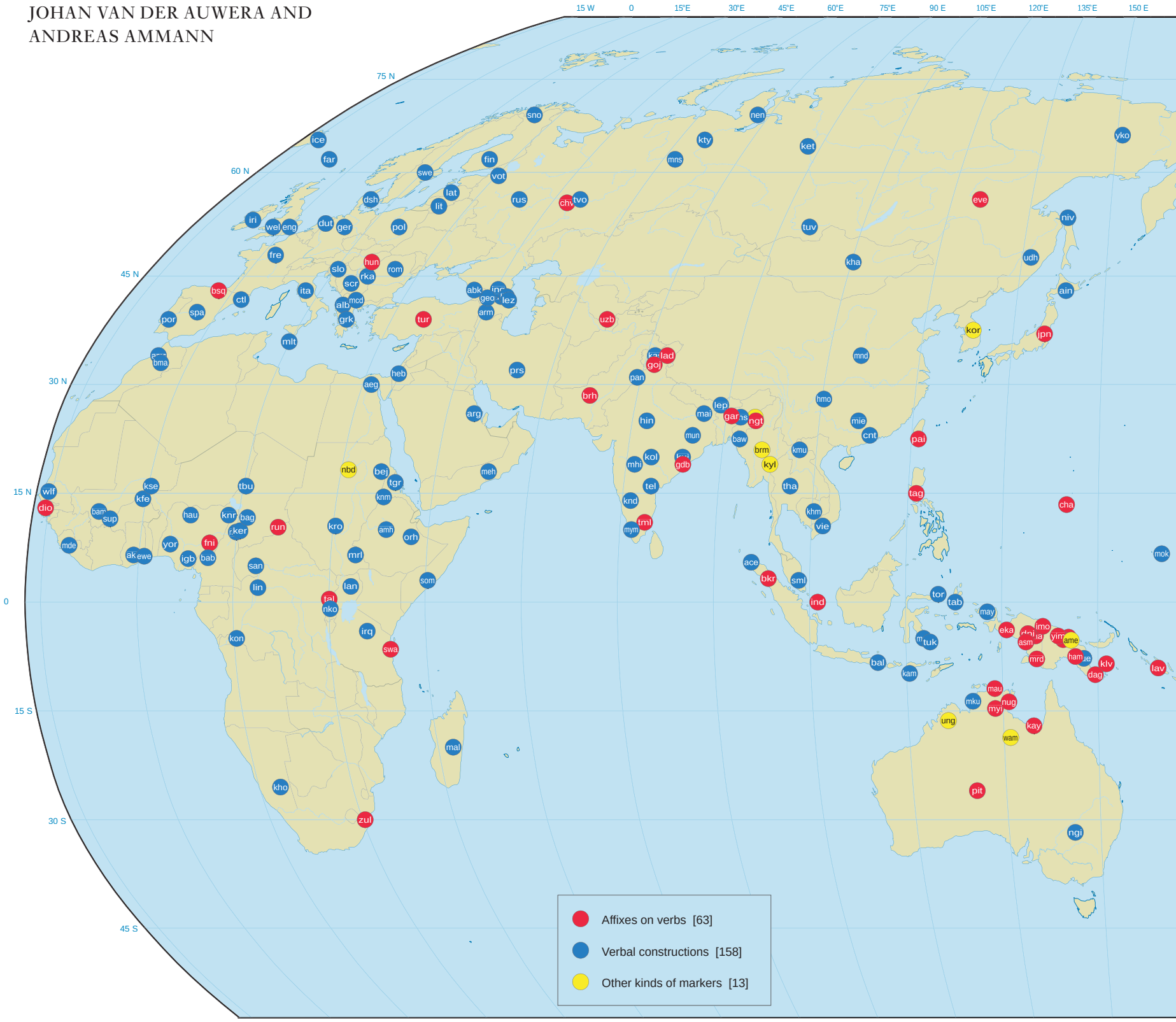
The most frequent type of language allows the expression of situational possibility with the use of (more or less) dedicated verbal constructions but not affixes. This type is particularly common in Europe, Africa, South Asia, and South-East Asia. It is rare only in New Guinea and in Australia.

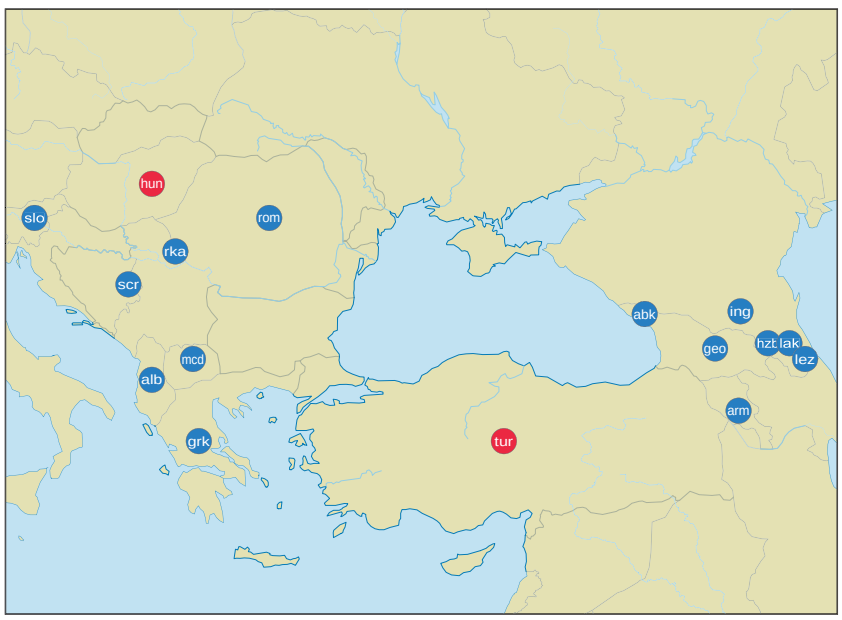
3 Theoretical issues

Situational possibility and, more generally, situational modality have received considerable attention in the study of European languages, in particular, because the markers employed in these languages are multifunctional and they have properties that clearly single them out as a special subclass of verbs, i.e. auxiliary verbs. English is one of these languages, and because English is one of the best-studied languages, the literature on the modals of English is voluminous. The multifunctionality of the English and English-like modals has been of interest for both semanticists (e.g. Palmer 1979, Coates 1983) and historical linguists (e.g. Traugott 1989), and the latter nowadays usually approach the topic from a grammaticalization point of view. Within grammaticalization theory typological work has been done by Bybee et al. (1994) and by van der Auwera and Plungian (1998). The status of auxiliaries has been approached within that framework as well (Heine 1993), but it has also received much attention from formal grammarians, from Ross (1969) onwards.

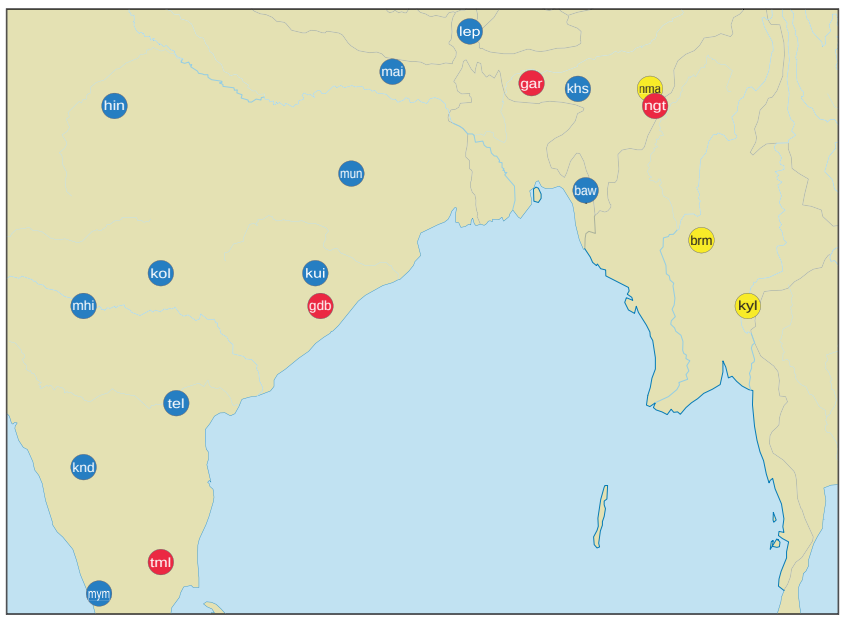
74 Situational Possibility

JOHAN VAN DER AUWERA AND
ANDREAS AMMANN

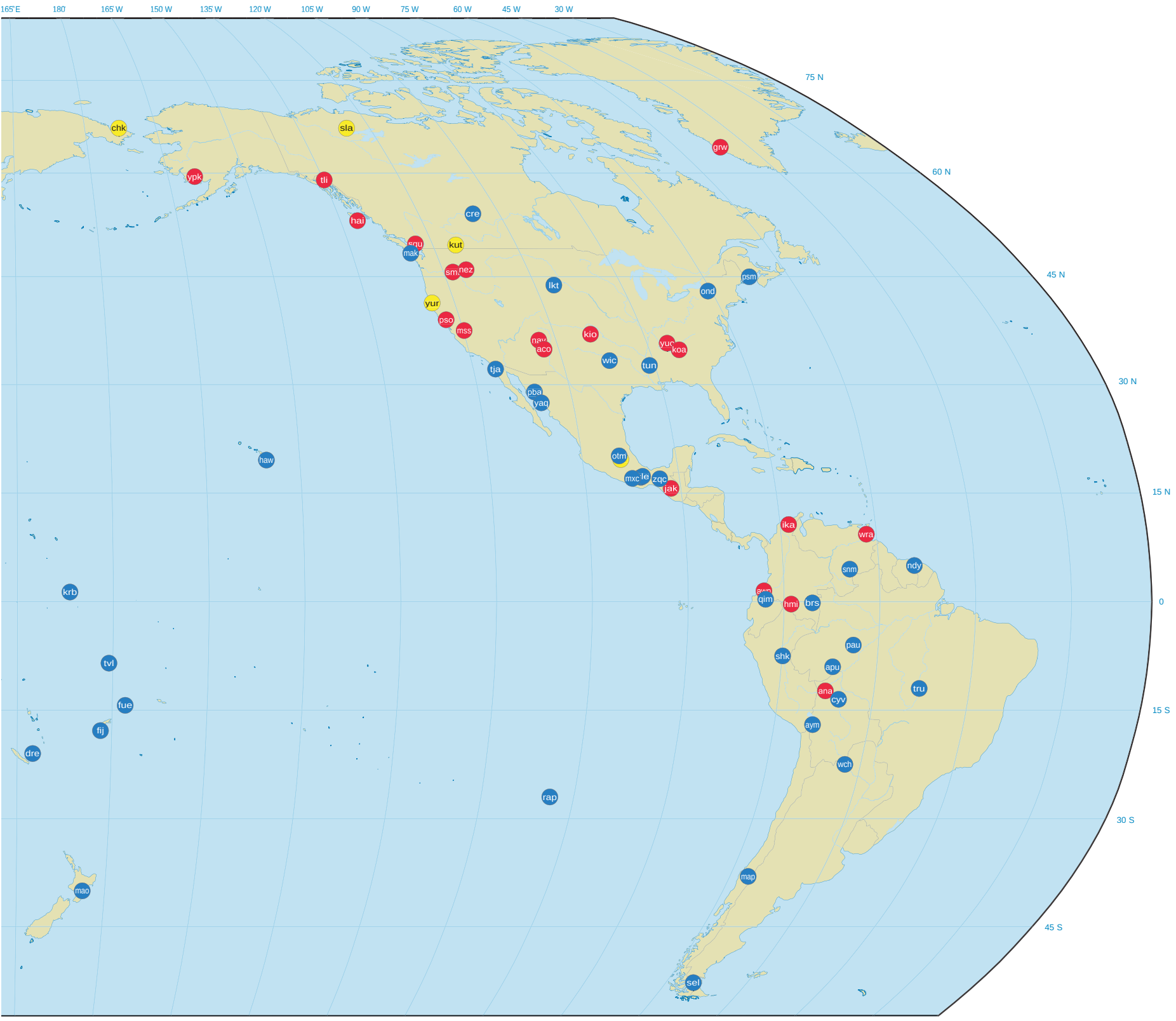


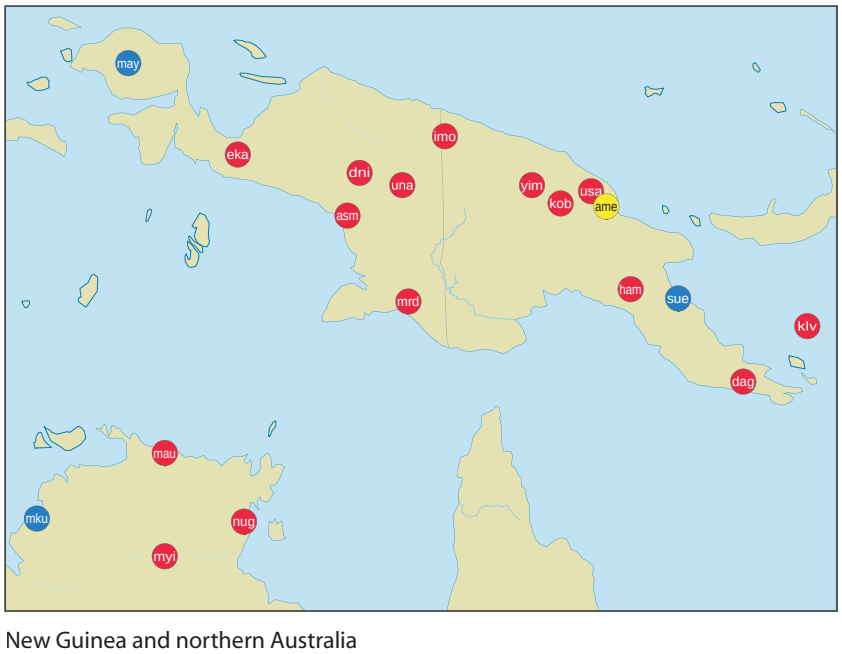


Balkans, Black Sea area, and Caucasus

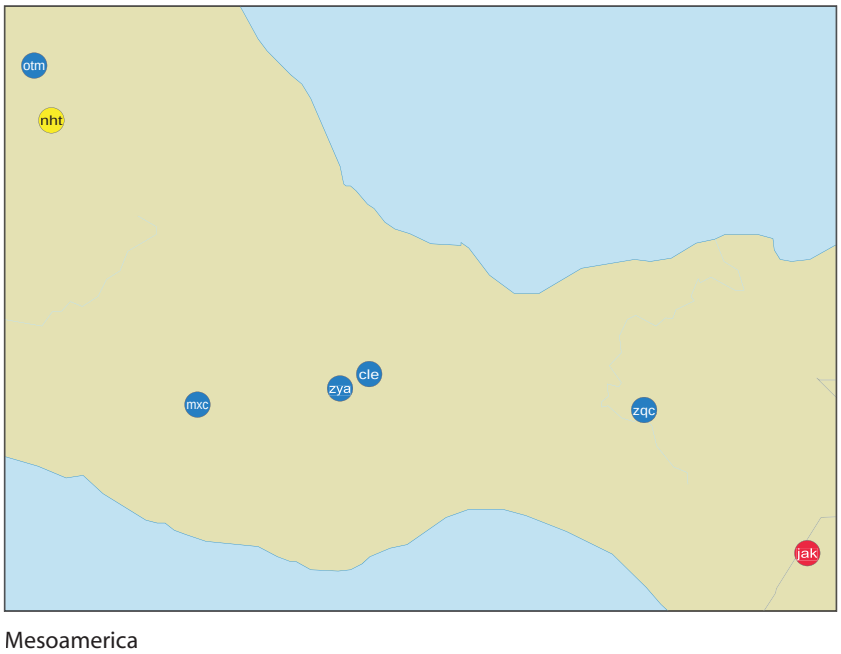


India, Bangladesh, and Myanmar





New Guinea and northern Australia



Mesoamerica

75 Epistemic Possibility

JOHAN VAN DER AUWERA AND ANDREAS AMMANN

1 Definition of values

This chapter deals with the kind of modality expressed by English *may* in (1).

(1) *John **may** have arrived.*

In (1) *may* indicates that the speaker holds that the proposition that John has arrived is not certain, relative to what he knows or to available evidence. It contrasts with the statement in (2).

(2) *John **must** have arrived.*

In (2) the speaker indicates that the proposition that John has arrived is certain, again relative to what he knows or to available evidence. This type of modality is commonly called **epistemic**, and it is distinguished from what we call **situational** modality, illustrated in (3) (see Chapter 74).

(3) a. *You **may** leave the hospital after a week.*
b. *You **must** leave the hospital after a week.*

In (3), *may* and *must* express a permission and an obligation. What is at issue is not any degree of certainty or uncertainty, but rather the degree to which the speaker tries to make a situation possible or necessary.

The map gives some indication of the strategies used to express epistemic possibility in positive main clauses. Three types of languages are distinguished. The definitions refer to three types of marking: verbal constructions, affixes on verbs, and other constructions, ordered in terms of increasing cross-linguistic frequency. Languages of the first type use the least frequent markers, whether or not they also use the more frequent ones. Languages of the second type use the medium frequency markers, whether or not they also use the most frequent ones. Languages of the third type only use the most frequent markers.

1. The language can express epistemic poss bility with verbal constructions.	65
2. The language does not express epistemic possibility with verbal constructions, but with affixes on verbs.	84
3. The language does not express epistemic possibility with verbal constructions or with affixes on verbs, but with other kinds of markers.	91
total	240

The first type shown is that of languages in which **verbal constructions have epistemic possibility as one of their meanings**. Example (4) illustrates this with Harar Oromo (Cushitic; Ethiopia). The verbal construction involves the verb *taha*, which also means 'be' or 'become'.

(4) Harar Oromo (Owens 1985: 78)
Ní *d'uf-t-i* ***taa-t-i.***
FOC come-F-IMPF become-F-IMPF
'She may come.'

As (4) shows, a verbal construction need not be uniquely dedicated to epistemic possibility. The verb responsible for the epistemic sense can also mean 'become' or 'be'. In Bambara (Mande; Mali), for example, epistemic possibility is rendered by a fixed expression *í b'ă soro*, which contains imperfective marking, a demonstrative pronoun, and the verb *soro* 'get'. The construction as a whole could be glossed as 'you'll get this', followed by the proposition (Valentin Vydrine, p. c.).

In some languages, a verbal construction may occur in an impersonal use, i.e. with a third person (singular). In Harar Oromo, an alternative to (4) involves the same verb in an invariable third singular masculine form.

(5) Harar Oromo (Owens 1985: 78)
Ní *d'uf-t-i* ***taha.***
FOC come-F-IMPV become.3SG M
'She may come.'

As illustrated with Harar Oromo *taha* and Bambara *soro*, epistemic modal verb constructions may contain verbs with a variety of literal or etymological meanings, including 'be/become' (Harar Oromo) and 'get' (Bambara). Verbs of propositional attitude like 'think' or 'suppose', however—i.e. expressions meaning 'I think' or 'I suppose', which are not grammaticalized and are probably universal—will not be included.

A language of the first type may have markers other than verbal constructions. Northern Saami (Uralic; Finland) can express epistemic possibility both with the "potential" inflection on the verb—marked with the suffix *-ž/žž/čč*—and with a verb *dáidit* for 'may, might'.

(6) Northern Saami (Sammallahti 1998: 77; Bartens 1989: 296)
a. *Bōđe-**žan**.*
come-POS.PRES.1SG
'I may come.'
b. *Mii* ***dáid-it*** *orrut* *dás* *geasi.*
we may-IND.PRES.1PL stay here.LOC summer.ILL
'We may be staying here all summer.'

The construction with the modal verb occurs more frequently than the "potential" inflection (Bartens 1989: 295–6), but this is irrelevant for our purposes. Northern Saami is thus assigned the value that specifies it as a language possessing epistemically modal verb constructions.

The second type includes **languages that cannot express epistemic possibility with verbal constructions, but that employ affixes on verbs**. This type is illustrated by Koasati (Muskogean; southeastern United States), which can employ the verbal suffix *-sahá:wa* with the meaning 'might be'.

(7) Koasati (Kimball 1991: 200)
Á:t-ok *hó:pa:-**sahá:w**-ok*
person-NOM.FOC be.sick-POS-SS.FOC
oh-hí:c-á:hi-k *amá:k.*
go-see-INTENT-SS go.PL.IMP
'Someone might be sick, go over and look, you all!'

Languages with clitics attached to verbs will also be included here. This is not to say that clitics and affixes are the same, but in cases where the clitic is described as occurring within a grammaticalized unit containing the verb, the similarity is rather high. For example, in Lakhota (Siouan; Nebraska and Minnesota), clitics—among them the enclitic *séčA/načhéčA* 'probably, maybe, I guess, I suppose'—occur in a fixed postverbal order and are frequently not stressed (Rood and Taylor 1996: 473–4).

Not all epistemic possibility affixes are uniquely dedicated to this function. The affix may also code related meanings like future —in Epena Pedee (Choco; Colombia; Harms 1994: 101)—and epistemic necessity—in Evenki (Tungusic; Siberia; Nedjalkov 1997: 265, and p. c.)—or more general meanings like subjunctive —in Gooniyandi (Bunaban; Western Australia; McGregor 1990: 193).

The third type **does not have either verbal constructions or verbal affixes to express epistemic possibility, but there are other markers**. These include adjectives like 'uncertain', or

'possible', as well as particles and adverbs, such as English *maybe* or the element *yilama* in Wardaman (Yangmanic; Northern Territory, Australia).

(8) Wardaman (Merlan 1994: 299)
Yilama *ya-yinyja* *ma-yinyja* *wurrugu* *ngarlg-ba*.
maybe IRR.3SG-go HAB.3SG-go 3N.SG.DAT call-PCL
'Maybe he'll go, he always goes to call them.'

As to clitics, we have already stated that clitics which occur on the verb and form a kind of unit with it are included in the affix type. But there are other kinds of clitics. For example, epistemic possibility is expressed in Jakaltek (Mayan; Guatemala) by a clitic *=m(i)*, which is usually attached to the first element of the sentence and is followed by the verb (see also Craig 1977: 86–7).

(9) Jakaltek (Day 1973: 55)
*Chin=**mi*** *wayoj*.
FUT.1SG=POS sleep
'I may be sleeping.'

Just as with modal affixes and modal verbs, these markers need not be dedicated expressions of (epistemic) possibility. One that is not is the phrase *yíghēe mɔ́* 'sometimes', which can be used for epistemic 'may' in Babungo (Bantoid; Cameroon; Schaub 1985: 229).

Note that epistemic particles and adverbs may have a verbal origin. In Welsh there is a marker *gallu* 'can, be able'. When used as a marker of situational possibility, *gallu* is a verb, but for epistemic possibility there is a form *efallai* 'perhaps'. It historically derives from *gallu*, but according to Fife (1990: 279), speakers no longer make the connection, and *efallai* is treated as an adverb by King (1993: 261). Interestingly, *gallu* does allow epistemic uses in relative clauses (Jones and Thomas 1977: 100), but since the present map only documents the expression of epistemic possibility in positive main clauses, these uses do not count here. On the basis of the adverb *efallai*, and the absence of a verbal construction (in positive main clauses) or any verbal affix, Welsh is coded as language of the "other" type.

2 Geographical distribution

In most European languages epistemic modality can be encoded by means of verbal constructions. This holds for most Indo-European and Uralic languages. In northern Eurasia, only Russian, Tuvan (Turkic; Altai region, Russia), and Mandarin show verbal constructions, while the other languages are divided between verbal affixes and the "other" type. In South Asia, the centre and south of India contrast as a verbal-auxiliary area with a verbal-affix area to the north-west and a mixed area to the north-east. Except for a relative sparseness of languages with epistemic verbal constructions, no pattern emerges in East Asia. The verbal construction type appears to be almost unattested in New Guinea and Australia. The languages of New Guinea prefer affixes, and for the languages of the Pacific our sample documents only markers other than verbal constructions or verbal affixes. On the African continent, the "other" type dominates in the centre. Semitic languages predominantly have verbal constructions, as do, apparently, the languages of West Africa. In the Americas not a single language investigated uses verbal constructions to express epistemic possibility.

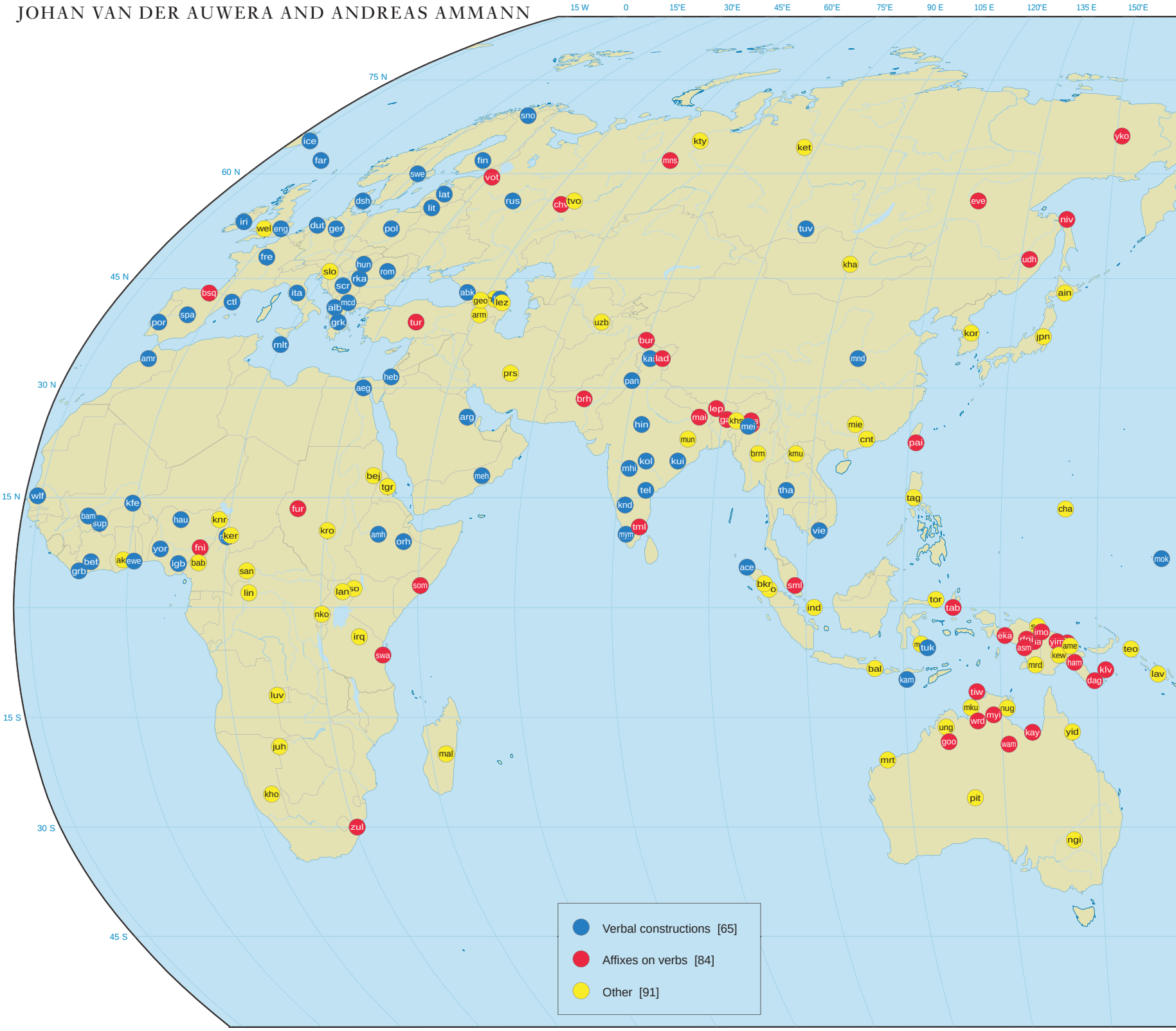
3 Theoretical issues

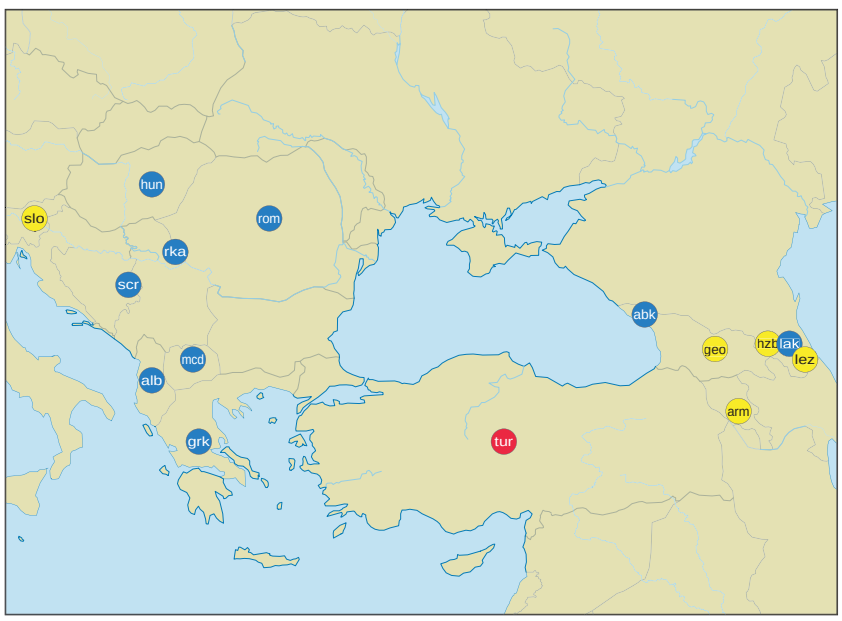
In European languages, most epistemic modality markers are historically related to situational ones. This has given rise to an ongoing debate about the conceptual relationship of the two. Cross-linguistic work on this issue is presented by Bybee et al. (1994) and by van der Auwera and Plungian (1998). Both stress that not all epistemic markers are related to situational ones.

A controversial point that needs to be noted is that there is little agreement as to what to include or not to include under the term "epistemic". The relationship with "evidential" marking is especially difficult to characterize (see Chapter 77). Are these two notions distinct, do they overlap, are evidentials a subtype of epistemic modality, or is it the other way around? We have taken a fairly liberal approach and included markers, for instance in Aymara (Aymaran; Peru and Bolivia) and Barasano (Tucanoan; Colombia), which occur in evidential-prominent systems but were described as having functions closely corresponding to epistemic possibility.

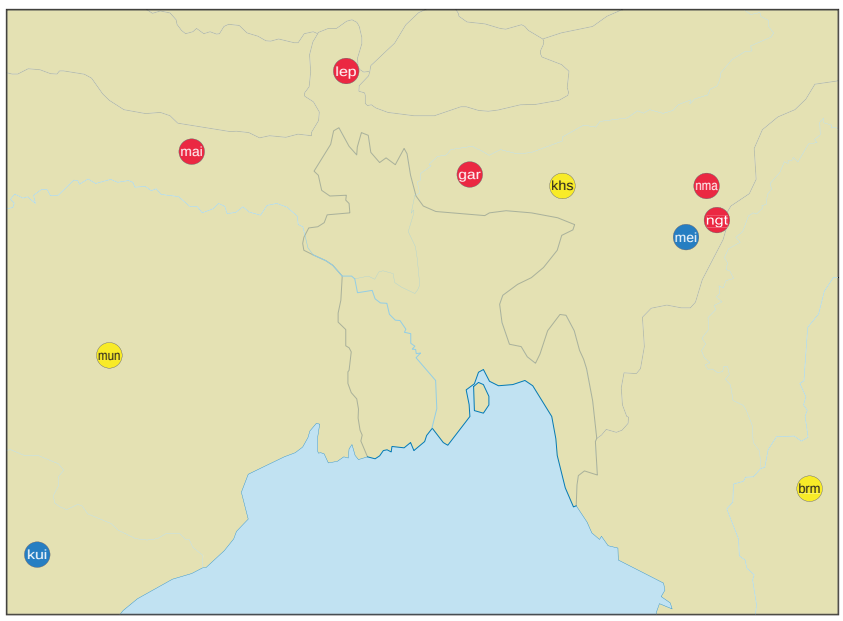
75 Epistemic Possibility

JOHAN VAN DER AUWERA AND ANDREAS AMMANN

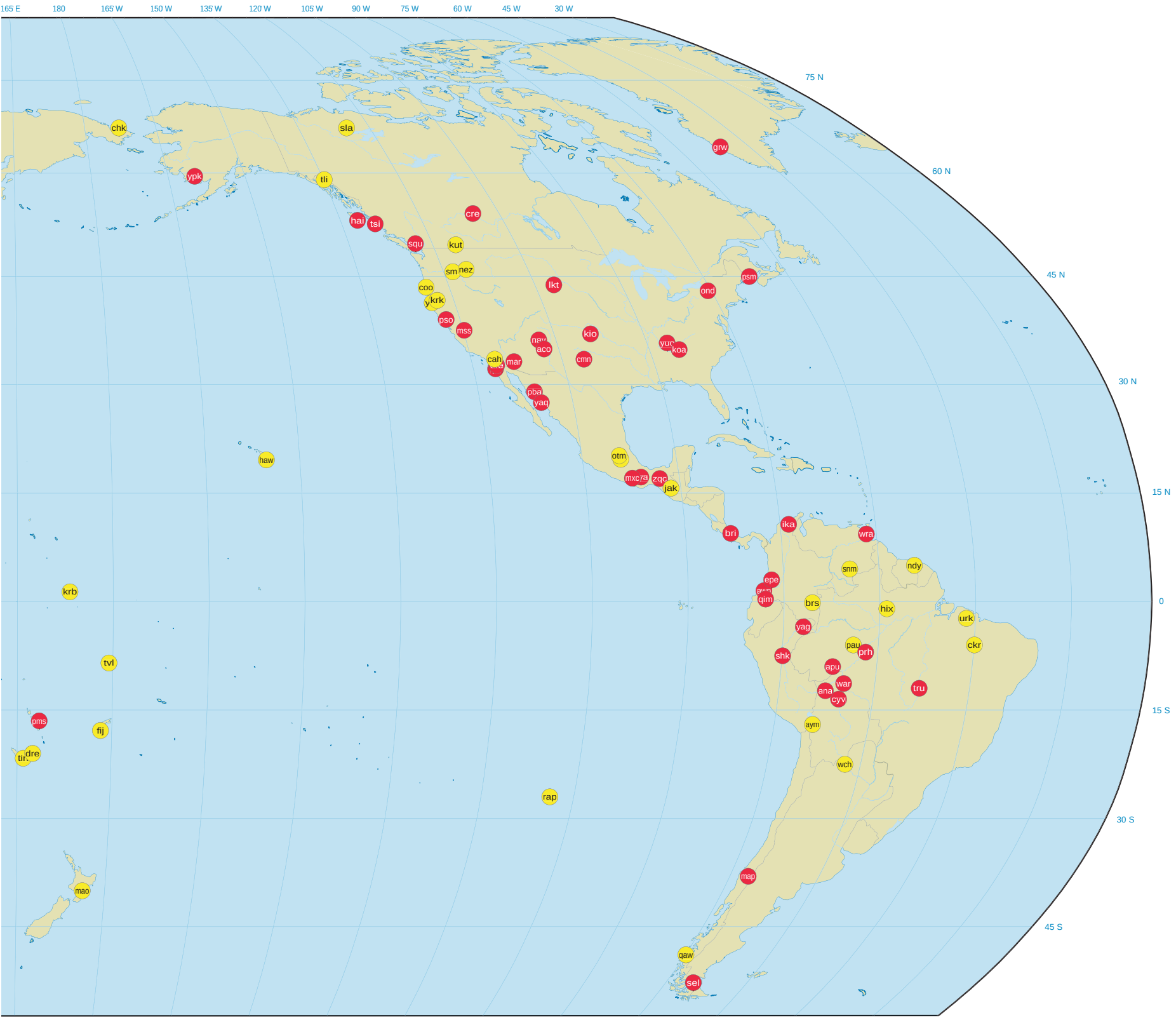


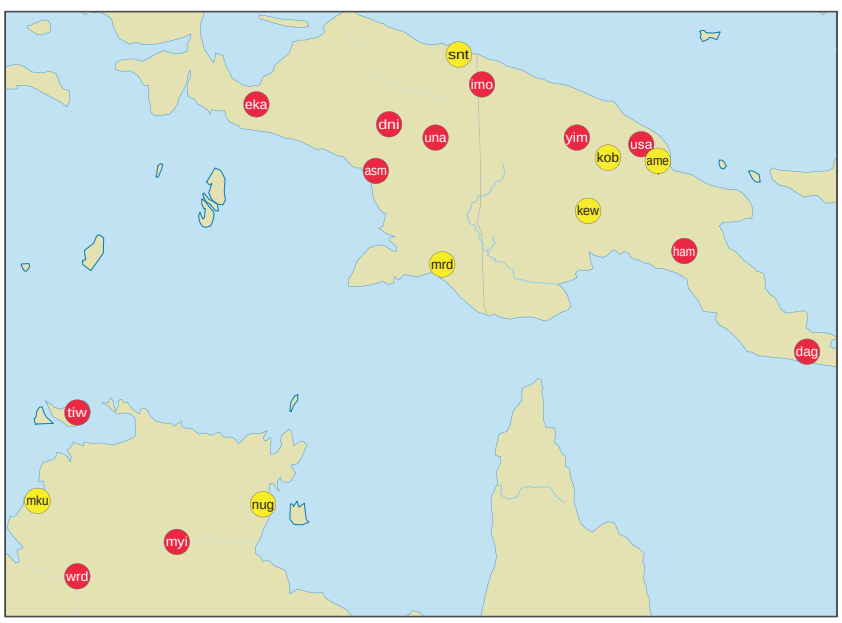


Balkans, Black Sea area, and Caucasus

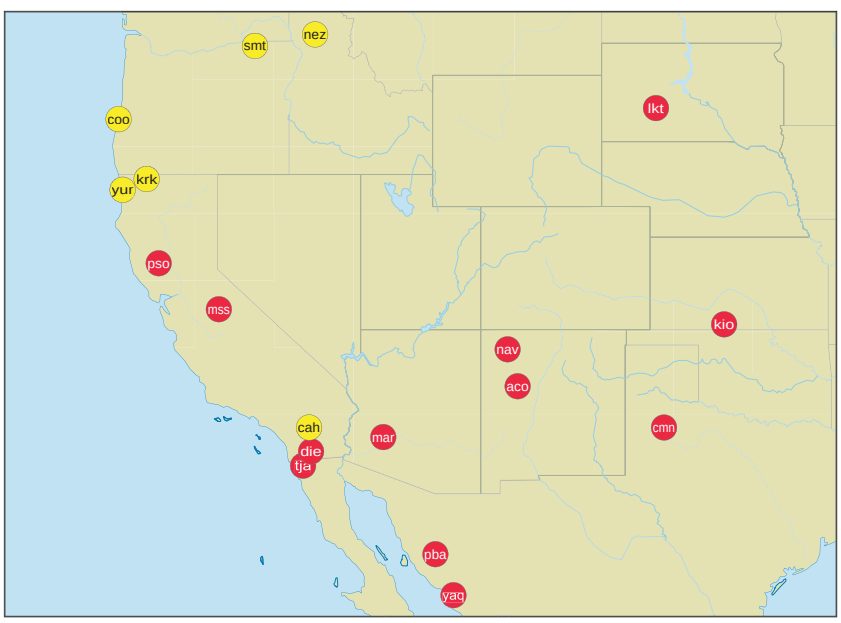


eastern India and Myanmar





New Guinea and northern Australia



western North America

76 Overlap between Situational and Epistemic Modal Marking

JOHAN VAN DER AUWERA AND ANDREAS AMMANN

1 Definition of values

Modality is a dimension of meaning that crucially involves notions of possibility and necessity. These may relate to circumstances that allow or necessitate a situation, as in (1). The permission in (1a) is a matter of the speaker allowing a situation, while with the obligation of (1b) the situation of the addressee's going home is deemed necessary.

(1) a. *You **may** go home now.*
b. *You **must** go home now.*

We call this type of modality **situational** (see Chapter 74).

Situational modality is distinguished from **epistemic** modality, as in (2a) and (2b) (see Chapter 75).

(2) a. *Bob **may** be mistaken about the cause of the accident.*
b. *Terry **must** be from Northumberland.*

In these instances of epistemic possibility (2a) and necessity (2b), the speaker asserts that a proposition is possibly or necessarily true, relative to some information or knowledge. If the proposition is only possibly true, the propositional attitude is that of uncertainty; if it is necessarily true, the propositional attitude is that of a high degree of certainty. Note that the English modals *may* and *must* can be used both for situational and for epistemic modality.

This chapter documents to what extent languages have identical markers for situational and epistemic modality. Three types will be distinguished.

1.	The language has markers that can code both situational and epistemic modality, both for possibility and for necessity.	36
2.	The language has markers that can code both situational and epistemic modality, but only for possibility or only for necessity.	66
3.	The language has no markers that can code both situational and epistemic modality.	105
	total	207

In the first type of language there are **markers that can express both situational and epistemic modality, for both possibility and necessity.** We will say that in this case the language shows **high overlap** between its situational and epistemic modal systems. Examples (1) and (2) have already shown English to be a case in point. Another example is West Greenlandic (Eskimo). The suffix *-ssa* 'should' has a situational necessity reading in (3a) and an epistemic necessity reading in (3b).

(3) West Greenlandic (Fortescue 1984: 292, 294)
a. *Inna-jaa-**ssa**-atit.*
go.to.bed-early-NEC-IND.2SG
'You must go to bed early.'
b. *København-mii-**ssa**-aq.*
Copenhagen-be.in-NEC-IND.3SG
'She must be in Copenhagen.'

In the same vein, the suffix *-sinnaa* has a situational possibility reading in (4a) and an epistemic possibility reading in (4b).

(4) West Greenlandic (Fortescue 1984: 293, and p. c.)
a. *Timmi-**sinnaa**-vuq.*
fly-can-IND.3SG
'It can fly.'
b. *Nuum-mut aalla-reer-**sinnaa**-galuar-poq . . .*
Nuuk-ALL leave-already-can-however-3SG.IND
'He may well have left for Nuuk already, but . . .'

In judging whether a situational and an epistemic marker are identical, we disregard syntactic differences. Modern Greek illustrates this case. For possibility the language uses the verb *boró* 'can'. This verb is inflected for person when used for situational modality, as in (5a), but appears in an impersonal construction as *borí* '(it) may (be)' in its epistemic function, as in (5b) (see also Joseph and Philippaki-Warburton 1989: 184).

(5) Modern Greek (Palmer 1986: 19–20, 29)
a. *Ta peðjá **borún***
the children can.IND.IMPF.PRES.3PL
na fíɣun ávrio.
that leave.IND.PFV.PRES.3PL tomorrow
'The children can leave tomorrow.'
b. *Ta peðjá **borí***
the children can.IND.IMPF.PRES.3SG
na fíɣun ávrio.
that leave.IND.PFV.PRES.3PL tomorrow
'Perhaps the children leave tomorrow.'

An identical set of verbs with the possibility of syntactic disambiguation is also found in Irish (Ó Siadhail 1989: 289), where modal verbs select a verbal noun in the situational function, but a finite complement in the epistemic function. In Tagalog (Austronesian; Philippines), speakers can choose a word order which is exclusively epistemic (Schachter and Otanes 1983: 273). Nevertheless, the markers remain the same, so that both Irish and Tagalog count as languages with highly overlapping modal systems.

In the second type of language, there are **markers for both situational and epistemic modality, but only for possibility or only for necessity, not for both.** We will say that such a language shows **some overlap** in its situational and epistemic modal systems. Ainu (isolate; Japan) has markers for situational and epistemic modality only in the case of necessity. The particle *kuni* 'must, should' is shown with a situational necessity use in (6a) and an epistemic necessity use in (6b).

(6) Ainu (Refsing 1986: 204; Tamura 2000: 118)
a. *A kor nispa, hokure **kuni** a*
I ATTR husband hurry NEC we
cisehe orun e hosipi.
house.of ALL you return
'My husband, you must hurry and return to our house.'
b. *Tapan hekaci poro yakun, isanispa*
this youth be.big if doctor
*ne an **kuni** p ne.*
as/into be NEC thing become
'When this child grows up, (s)he should become a doctor.'

There is no such overlap for possibility. The central marker of situational possibility (*askay*), for instance, does not appear to be used for epistemic functions.

(7) Ainu (Refsing 1986: 207–8.)
*Hure konno, a e **easkay.***
be.red when INDEF eat can
'When they have become red, one can eat them.'

A language in which an overlap between situational and epistemic modality occurs for possibility, but not for necessity, is Hungarian (Kenesei et al. 1998: 314–17). The relevant marker is the verb *lehet*, illustrated with a situational use in (8a) and an epistemic one in (8b).

(8) Hungarian (Kenesei et al. 1998: 315, 316)

a. *Haza* ***lehet*** *men-n-em?*
to.home may go-INF-1SG
'Am I allowed to go home?'

b. ***Lehet,*** *hogy* *jöv-ök.*
may that come-IND.PRES.INDEF.1SG
'I may be coming.'

The two uses are different syntactically: in its situational use *lehet* selects an infinitive, whereas in its epistemic use it selects a complement clause. But the same marker *lehet* occurs in both, and we thus count Hungarian as a language with some overlap.

In the third type of language, there are **no markers that code both situational and epistemic modality, neither for possibility nor for necessity**. The language will be said to show **no overlap** in its systems for situational and epistemic modality. Evenki (Tungusic; Siberia) is a case in point. It uses a modal participle *-d'AngA* (or a future tense) for situational possibility (9a) and a suffix *-mAchin* for situational necessity (9b). Then there are epistemic suffixes, viz. *-nA, -rkA*, and *-rgu*, which all cover both epistemic possibility (9c) and necessity (9d) and differ as to temporal-aspectual reference (present or recent past, past, and habitual respectively). The two sets of markers are fully distinct.

(9) Evenki (Nedjalkov 1997: 264, 265, 269, and p. c.)

a. *Ulguchen-**d'enge*** *bejetken* *amakan* *eme-d'e-n.*
tell-PTCP boy soon come-FUT-3SG
'The boy who will be able to tell will come soon.'

b. *Minggi* *girki-v* *ilan-duli*
my friend-1SG.POSS three-PROL
chas-tuli *suru-**mechin**-in.*
hour-PROL go.away-NEC-3SG
'My friend must go/leave in three hours.'

c. *Ga-**na**-m.*
take-POS-1SG
'Maybe I take/took recently.'

d. *Su* *tar* *asatkan-me* *sa:-**na**-s.*
you that girl-ACC.DEF know-NEC-2PL
'You probably know that girl.'

If a language that lacks specialized expressions for situational and/or epistemic modality makes use of the same general tense-aspect-mood marker in both subdomains, this is not considered to be an instance of the relevant overlap, as the primary function of the marker in question does not fall into either field. One such case is the particle *ki* in Yurok (Algic; California; Robins 1958: 99–100), which can stand for 'may', 'can', 'ought', and 'must', i.e. covering virtually all the territory of situational and epistemic modality, but also 'will'. In fact, of the various future markers in Yurok, this particle *ki* comes closest to being a purely temporal marker.

2 Geographical distribution

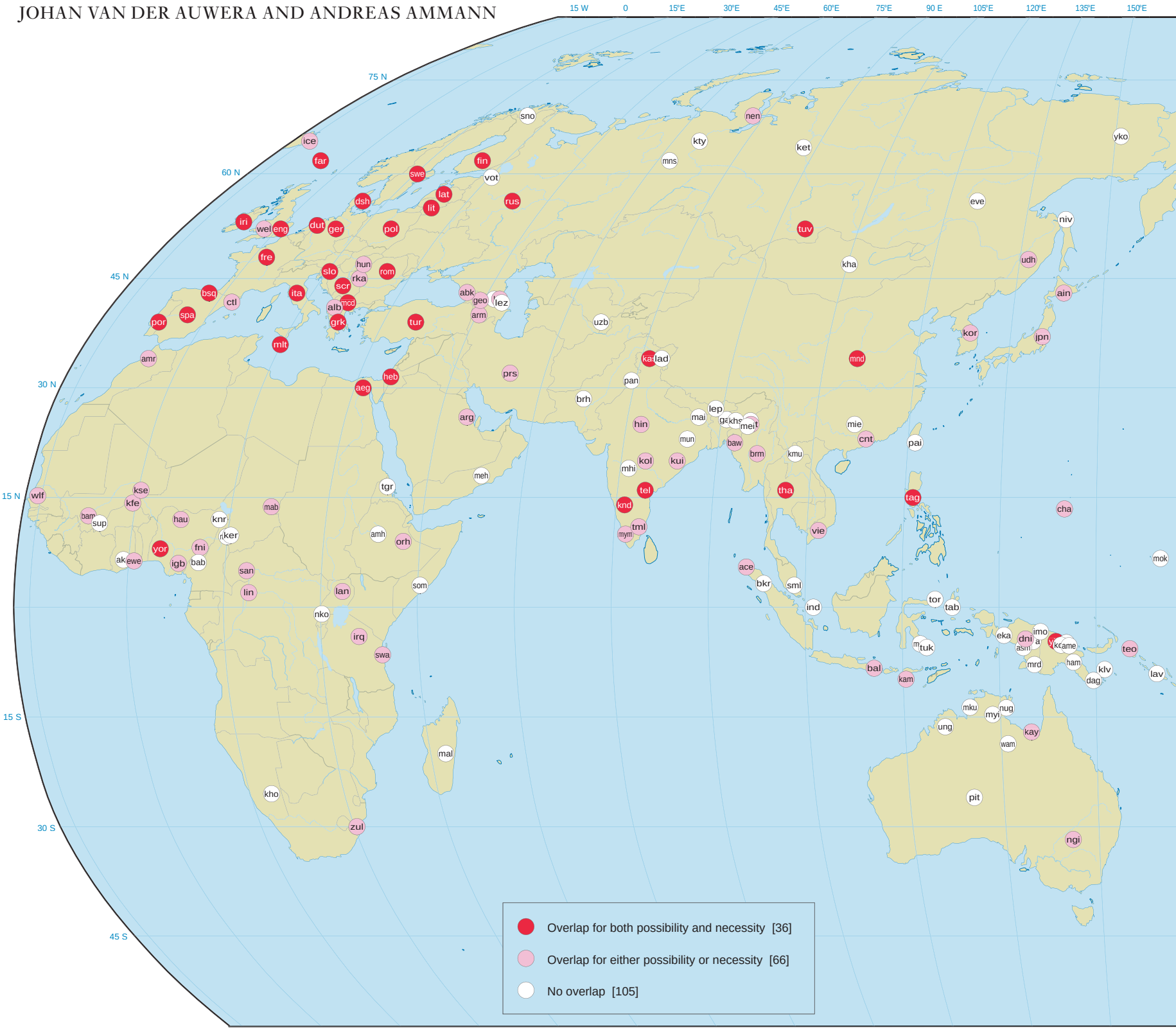
The first and most important finding is that high overlap is characteristic of Europe. Here, and only here, nearly all languages have highly overlapping systems, and those that do not still show at least some overlap. High overlap is also found in Turkish, Hebrew, and Egyptian Arabic. Outside of this area, languages without any overlap are in the majority. Languages with a high degree of overlap occur only sporadically in Asia, and we did not discover any examples in sub-Saharan Africa or in the Americas. An interesting case in between Europe and America is West Greenlandic. Like Yup'ik it expresses both situational and epistemic possibility with verbal morphology. Yup'ik, however, seems to have no overlapping modals. West Greenlandic does, for both possibility and necessity, and whereas the overlap for necessity is old, the one for possibility is recent and probably due to contact influence from Danish (Fortescue 1984: 293, and p. c.). An area in which absolutely no overlap was detected is, roughly speaking, central South America. North of the Panama Canal, the absence of overlap also predominates, and the languages with partial overlap do not cluster in any particular area.

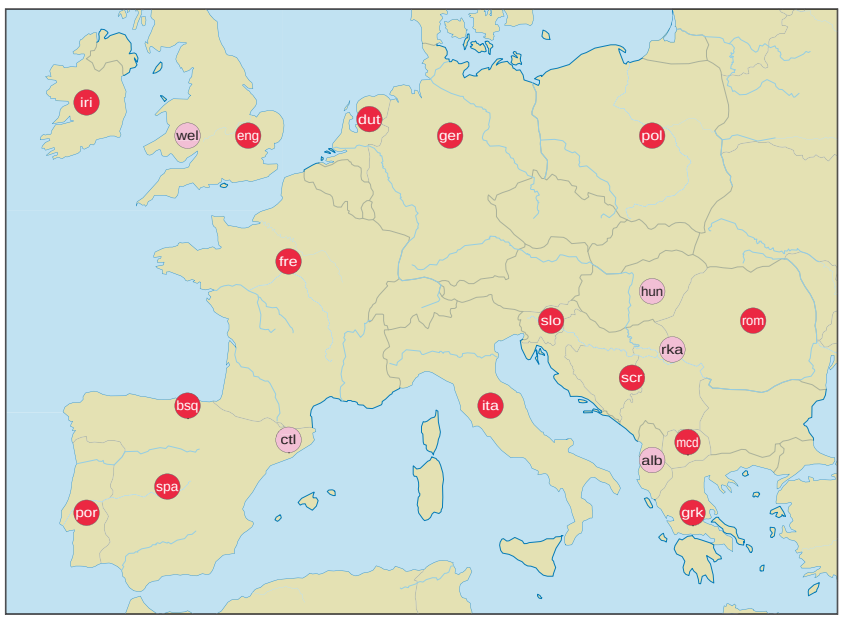
3 Theoretical issues

There are two ways of approaching the relationship between what we have called "situational" and "epistemic" modality. One is more synchronically oriented and focuses on the relationship of semantics and morphosyntactic expressions. Excellent work, for English, includes Coates (1983) and Perkins (1983). Such language-specific studies are relevant to the topic because it must first be determined how it can be established whether a modal conveys situational or epistemic meanings or both. The foundation of comparative research was laid by Palmer (2001 [1986]). Bybee et al. (1994) and van der Auwera and Plungian (1998) present a comparative and diachronic approach drawing largely on grammaticalization theory. The focus has been on determining paths of semantic change.

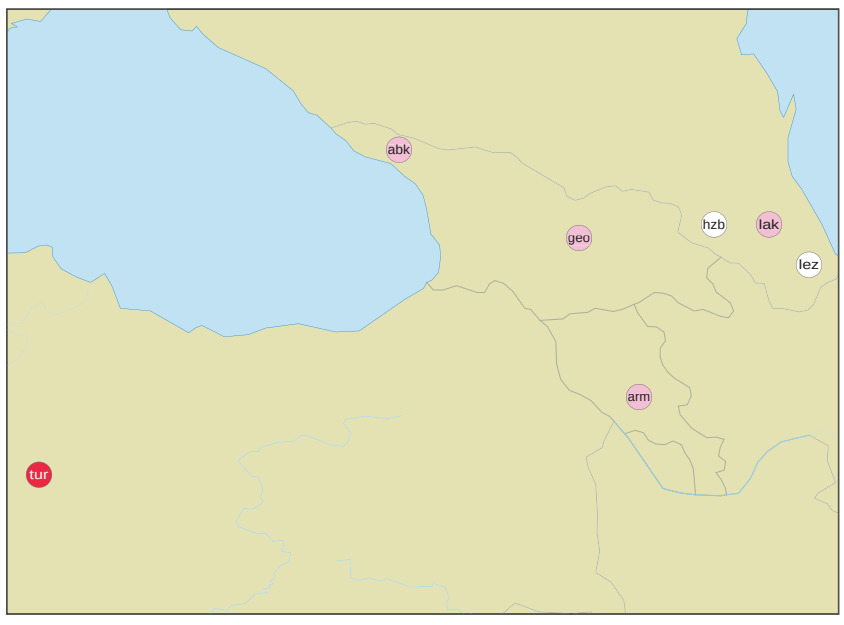
76 Overlap between Situational and Epistemic Modal Marking

JOHAN VAN DER AUWERA AND ANDREAS AMMANN

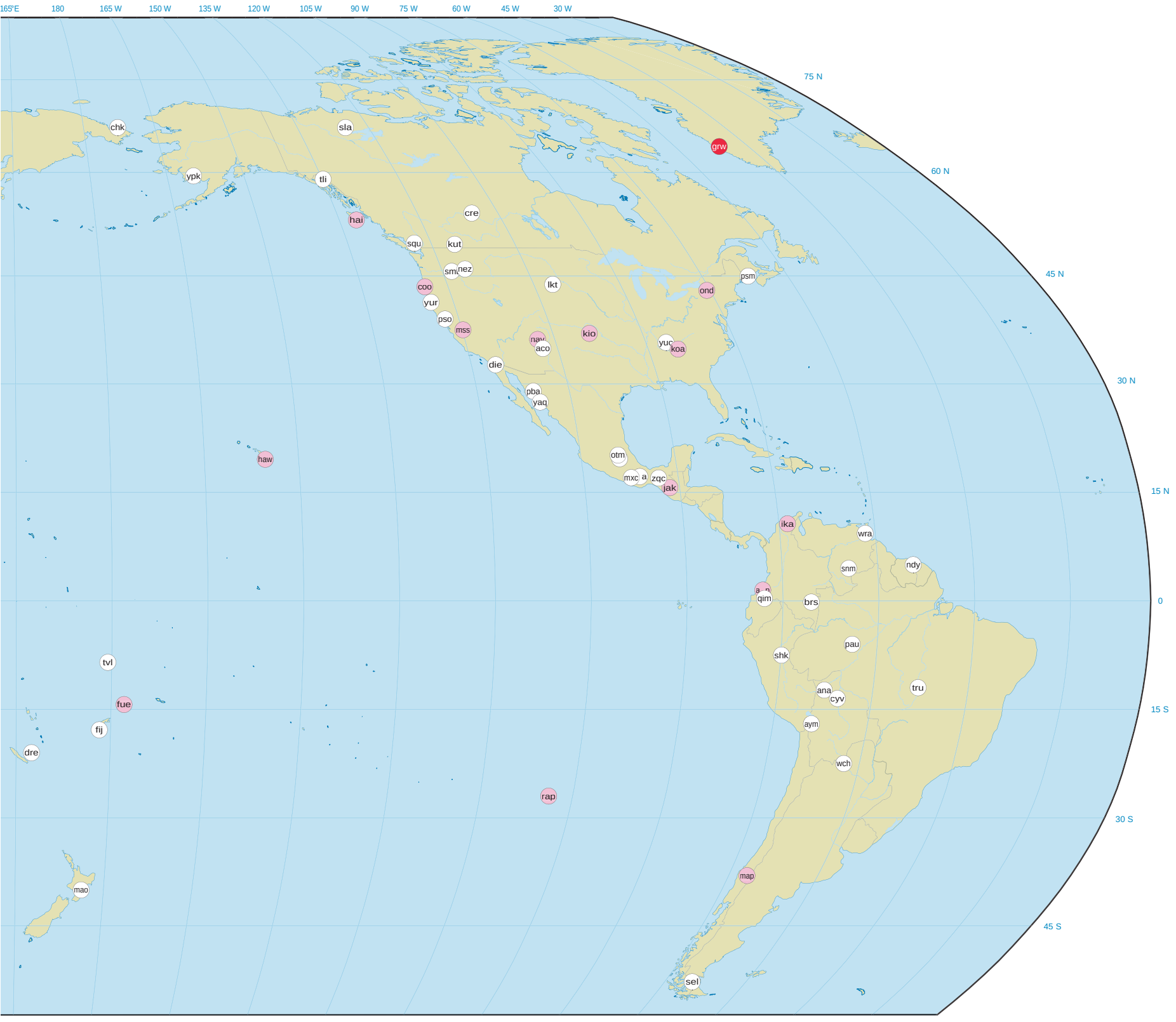


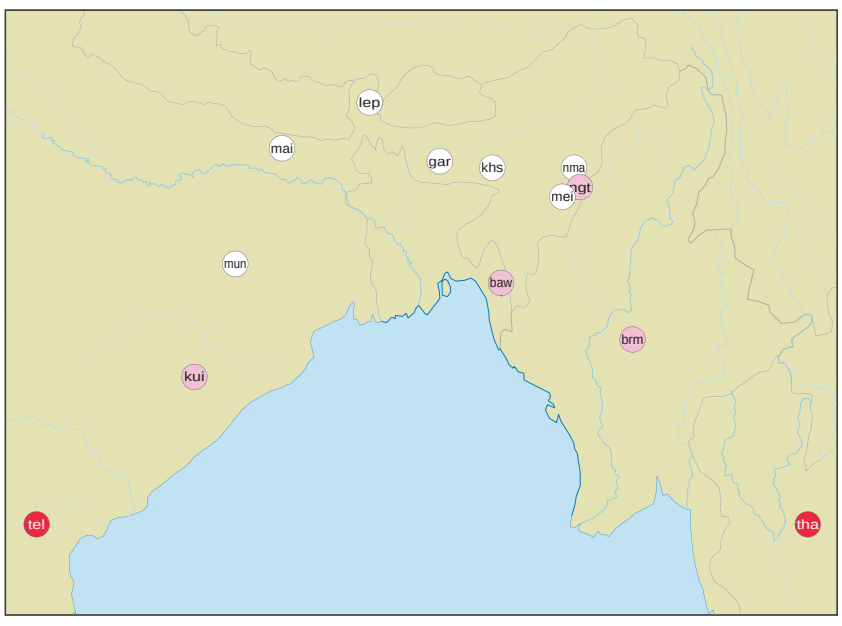


Europe

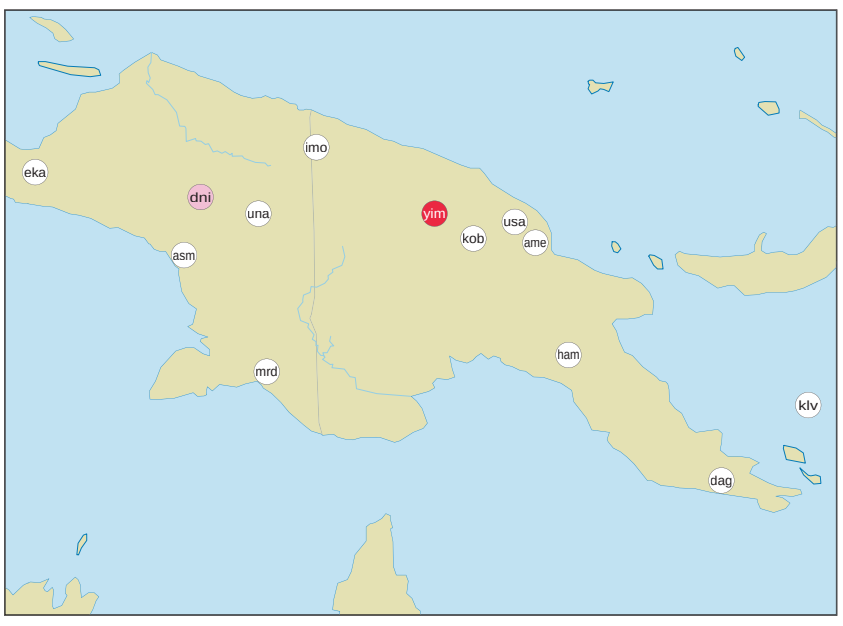


Caucasus and adjacent areas





from eastern India to Thailand



New Guinea

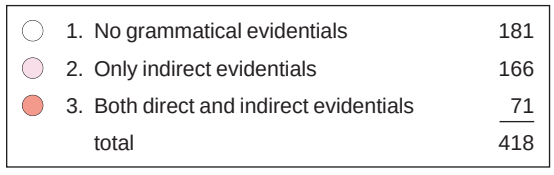
77 Semantic Distinctions of Evidentiality

FERDINAND DE HAAN

1 Defining the values

This map and the one following it show the presence of grammatical markers of evidentiality. On this map the semantic coding of evidentiality is represented, while Map 78 deals with the morphological coding of evidentiality.

On Map 77, the following feature values are shown:



Markers of evidentiality express the evidence a speaker has for his/her statement. For the purposes of this map and the following one, evidentiality is divided into two broad subcategories, direct evidence and indirect evidence. Only grammaticalized evidentials have been included in this study; this excludes lexical adverbs, such as *reportedly* in *John is reportedly ill*, and biclausal constructions, such as *It is said that John is ill*. Auxiliary verbs and particles have been included. Also, a distinction is made between evidentiality, which is used only for the evidence for the statement, and epistemic modality, which marks the speaker's degree of confidence in the statement (see Chapter 75). Although it is possible for evidentials to develop from epistemic modals (see Chapter 78), this is not common.

Direct evidentials are used when the speaker has some sort of sensory evidence for the action or event he/she is describing. Normally, a direct evidential denotes **visual evidence**. Example (1), from Fasu (Trans-New Guinea; Papua New Guinea; Loeweke and May 1980: 71), shows a visual evidential in the form of the circumfix *a-* . . . *-re* :

(1) Fasu
a-pe-re
VIS-come-VIS
'[I see] it coming.'

Other sensory information can be coded in evidentials as well. Example (2), from Koasati (Muskogean; Louisiana and Texas; Kimball 1991: 207), shows an **auditory evidential**, which expresses the idea that the action described, in this case the charring of meat, was perceived by hearing.

(2) Koasati
nipó-k *aksóhka-ha*
meat-SUBJ char-AUD
'It sounds like the meat is charring.'

The other senses are rarely expressed individually. If they are expressed with an evidential, they are usually combined with auditory evidence to form a **nonvisual sensory evidential**. Example (3), from Tuyuca (Eastern Tucanoan; Colombia and Brazil; Barnes 1984: 260), shows a nonvisual sensory evidential. In Tuyuca, evidential affixes also express tense (past and present, but not future), person, number, and gender. Example (3a) shows the auditory use of this evidential, while (3b) shows an example of one of the other senses (with a change in tense).

(3) Tuyuca
a. *mũtúru* *bisí-ti*
motor roar-NON.VIS.OTHER.PST
'The motor roared.' (I heard it)
b. *páaga* *punĩ-ga*
stomach hurt-NON.VIS.OTHER.PRES
'My stomach hurts.' (I feel it)

In the majority of cases, however, there is one general direct evidential, covering all modes of sensory evidence, although it is usually used most often for visual evidence. Example (4) comes from Kewa (Papua New Guinea; Franklin 1971: 50). In this example, the most common interpretation of *-ha* is that it denotes visual evidentiality, but other forms of direct evidence are possible.

(4) Kewa
íra-a-na
cook-3.SG.NEAR.PST-DIR.EVD
'He cooked it.'

Indirect evidentials are used when the speaker was not a witness to the event but learned of it after the fact. There are two broad subcategories, inference and quotative. **Inferential evidentials** are used when the speaker draws an inference on the basis of available physical evidence. Example (5) comes from Khalkha (Mongolian; Street 1963: 129):

(5) Khalkha
ter *irsen* *biz*
he come INFER
'He must have come.'

Quotatives (also known as *reportatives*, *hearsay*, or *second-hand evidentials*) are used when the speaker has been told about the action or event by another person. Example (6) is from Lezgian (Nakh-Daghestanian; eastern Caucasus; Haspelmath 1993: 148):

(6) Lezgian
Qe *sobranie* *že-da-lda.*
today meeting be-FUT-QUOT
'They say that there will be a meeting today.'

As is the case with direct evidentials, more often than not both types of indirect evidence are grouped together into one general indirect evidential. This happens for instance in Dutch, where the verb *moeten* 'must' (also used as an epistemic or deontic modal, like its English cognate verb *must*) can be used for unspecified indirect evidence:

(7) Dutch
Het moet een goede film zijn.
'It seems to be a good movie.' (I have no direct evidence)

Unlike the Dutch verb *moeten*, English *must* does not have an evidential sense. Whereas Dutch *moeten* in its evidential sense has no epistemic modal interpretation, English *must* does. The latter is therefore not an evidential in the way this term is used in this study (see de Haan 1999 for further details).

Many languages make a distinction of witnessed (direct) versus unwitnessed (indirect) actions in the past tense. These include Kartvelian languages like Georgian (Boeder 2000), many Turkic languages (see (8) below), as well as Komi-Zyrian (Finno-Ugric; Leinonen 2000), Haida (isolate; Alaska and British Columbia; Swanton 1911), and Ika (Chibchan; Colombia; Frank 1990). Example (8) is from Turkish (Aksu-Koç and Slobin 1986), where a distinction is made between witnessed past (the morpheme *-di*) and unwitnessed (*-miş*).

(8) Turkish
a. *Ahmet* *gel-di.*
Ahmet come-PST.DIR.EVD
'Ahmet came.' (witnessed by the speaker)

b. *Ahmet gel-miş.*
Ahmet come-PST.INDIR.EVD
'Ahmet came.' (unwitnessed by the speaker)

Because evidentials are used to describe the speaker's involvement with events, they tend to occur in realis contexts, especially in past-tense situations (cf. example (8) above). Nevertheless, evidentials do occur in what can be described as irrealis situations. Example (9) from Barasano (Eastern Tucanoan; Colombia; Jones and Jones 1991: 116) shows an interrogative evidential, and (10), from Tsova-Tush (Nakh-Daghestanian; Georgia; Holisky and Gagua 1994: 180), shows an evidential with a future event.

(9) Barasano
rase *yã-ro-hari* *ĩ*
toucan be-PRES.DIR-Q 3SG.M
'Is he a toucan?'

(10) Tsova-Tush
tit'-o-ra-lo
cut.FUT-FORMANT-PST-QUOT
'S/he will be cutting it apparently.'

There are no languages in the sample with grammaticalized direct evidentials, but no grammaticalized indirect evidentials.

2 Geographical distribution

Evidentials are common in the languages of the world, more common than usually assumed. They are found on every continent, but some very clear areal patterns emerge from the map: (1) evidentials are almost completely absent in Africa; (2) languages having only indirect evidentials are common in Europe; (3) languages with both direct and indirect evidentials appear to be clustered in the western United States, the western Amazon region, the Caucasus, and in the Himalayan and adjacent areas, but they are by no means restricted to these areas; (4) languages of the Americas are very likely to have at least indirect evidentials; (5) languages of the Pacific area, including New Guinea, are slightly more likely to have no evidentials. As a general statement, evidentiality would appear to be more of an areal feature than a genetic feature.

The most remarkable feature of the map is the almost complete absence of evidentials in Africa. The only African languages with evidentials in the sample are Afrikaans, a Germanic language spoken in South Africa, Lega, a Bantu language in central Africa (which has both a direct and an indirect evidential according to Botne 1995), and Beja, a Cushitic language with a quotative construction.

The languages with indirect evidentials in Europe are mostly Germanic (though not English). Romance languages do not seem to have evidentials, with the exception of French, probably under the influence of Germanic. The same can be said for Finnish, which has evidentials derived from modal verbs, as do the Germanic languages. Evidentiality can possibly be considered to be part of the Balkan Sprachbund, even though languages such as Modern Greek and Romanian do not seem to have them. The presence of evidentiality in Bulgarian, Macedonian, and Albanian may be due to Turkish influence.

In Asia, there are two areas with a high concentration of evidential systems: the Caucasus and the Himalayan area. In the languages of the Caucasus, especially the Kartvelian group, evidentiality is encoded with tense and aspect morphemes, although it can also be expressed with separate suffixes. On the South Asian subcontinent, the Indo-Aryan languages in general lack evidentials, except those that have been in contact with Tibeto-Burman or Iranian languages. In the southern part of the subcontinent, the Dravidian languages do have evidentials. The Tibeto-Burman family is rich in evidentials, although the phenomenon would seem to occur more frequently in the Tibetic group than elsewhere. Mon-Khmer languages and Tai languages do not seem to have evidentials.

In the Pacific region, evidentials are not found in the Polynesian languages or outside the Australia/New Guinea area, but otherwise are sporadically documented.

Evidentials are widespread in the Americas. In North America, the largest evidential systems can be found in California (especially the Pomoan languages) and the south-west (in Western Apache and the Uto-Aztecan languages, but also in Kiowa-Tanoan). However, evidentials occur in almost every language family in the Americas. In South America, the largest evidential systems occur in the Vaupés River region in Western Brazil and adjacent areas.

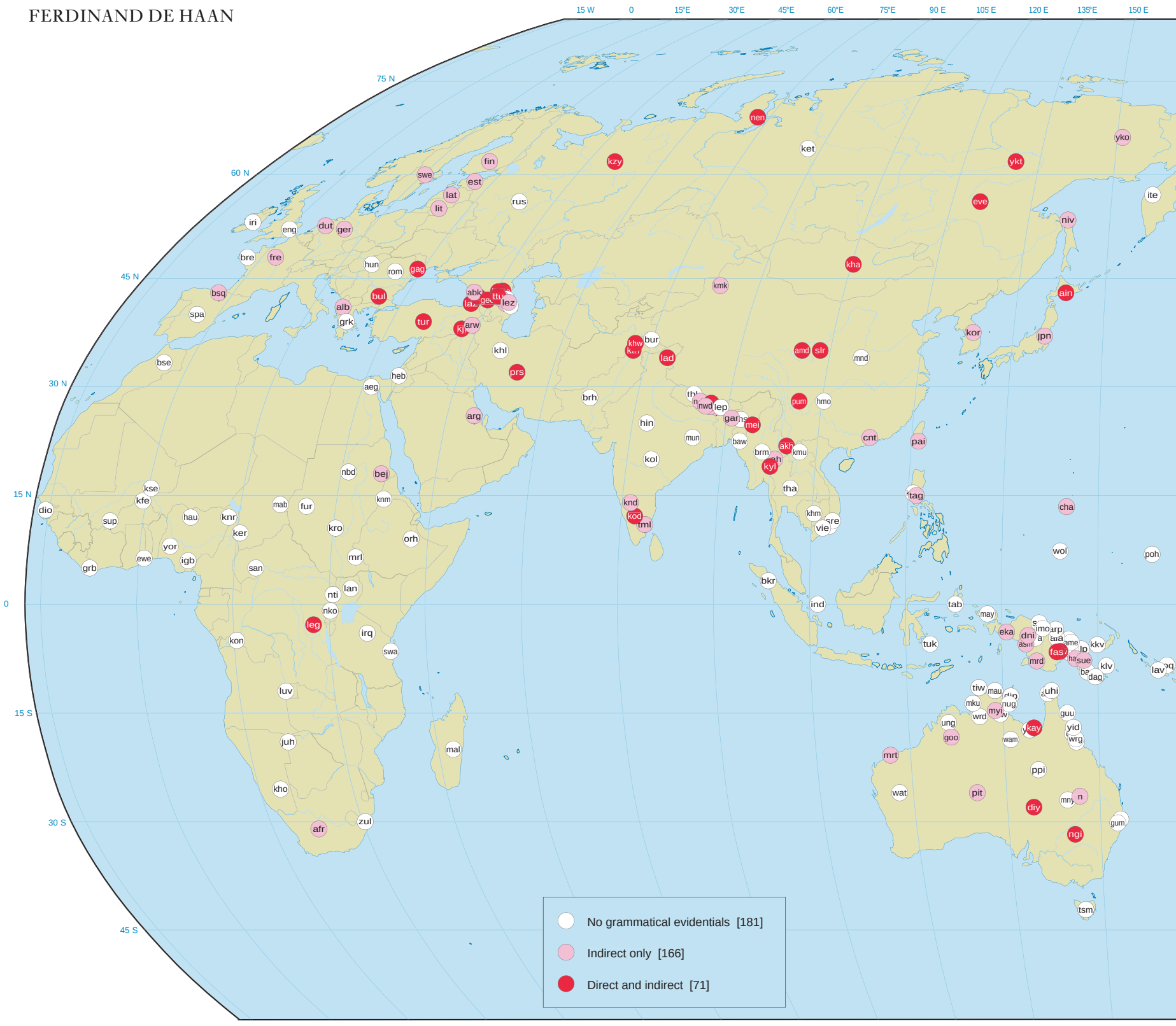
3 Theoretical issues

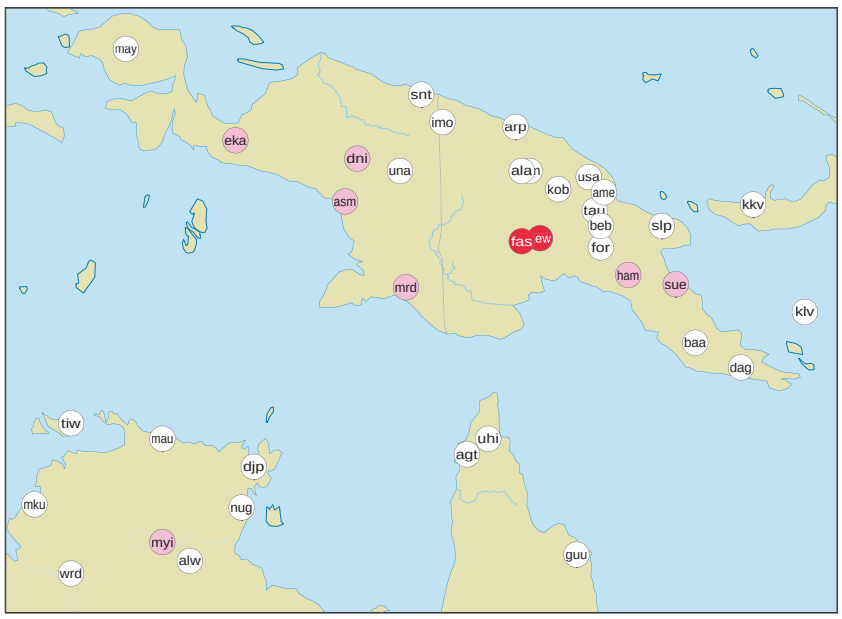
Current research in evidentiality focuses on the interaction of evidentiality with other categories, such as epistemic modality (Palmer 1986, de Haan 1999), tense, and deixis (Floyd 1999, de Haan 2001). Also of interest is the role of evidentiality as a possible areal feature, given the ease with which it spreads from language to language (see Haarmann 1970, the papers in Johanson and Utas 2000, and Aikhenvald and Dixon 1998).

Further recent contributions to the cross-linguistic study of evidentiality are Guentchéva 1996, Aikhenvald and Dixon 2003, and Aikhenvald 2005.

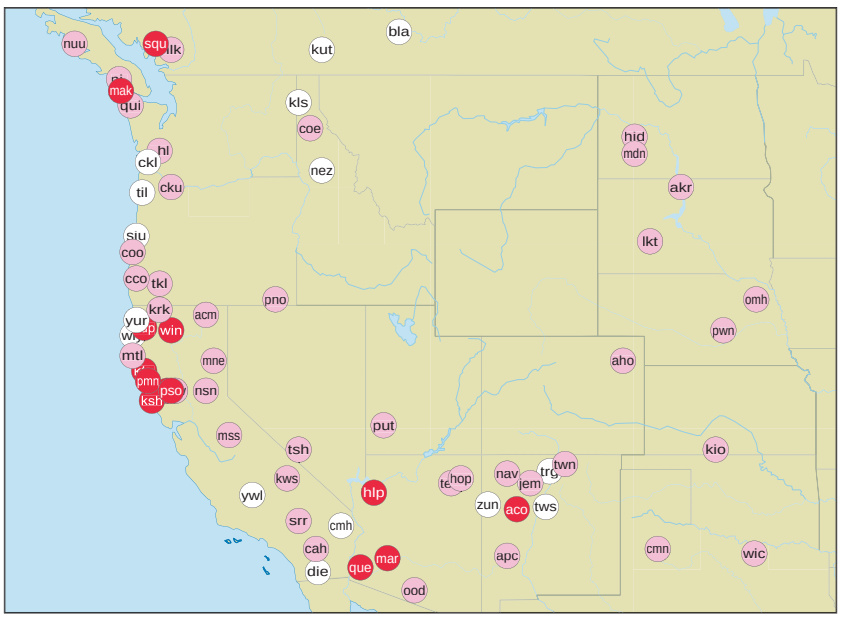
77 Semantic Distinctions of Evidentiality

FERDINAND DE HAAN

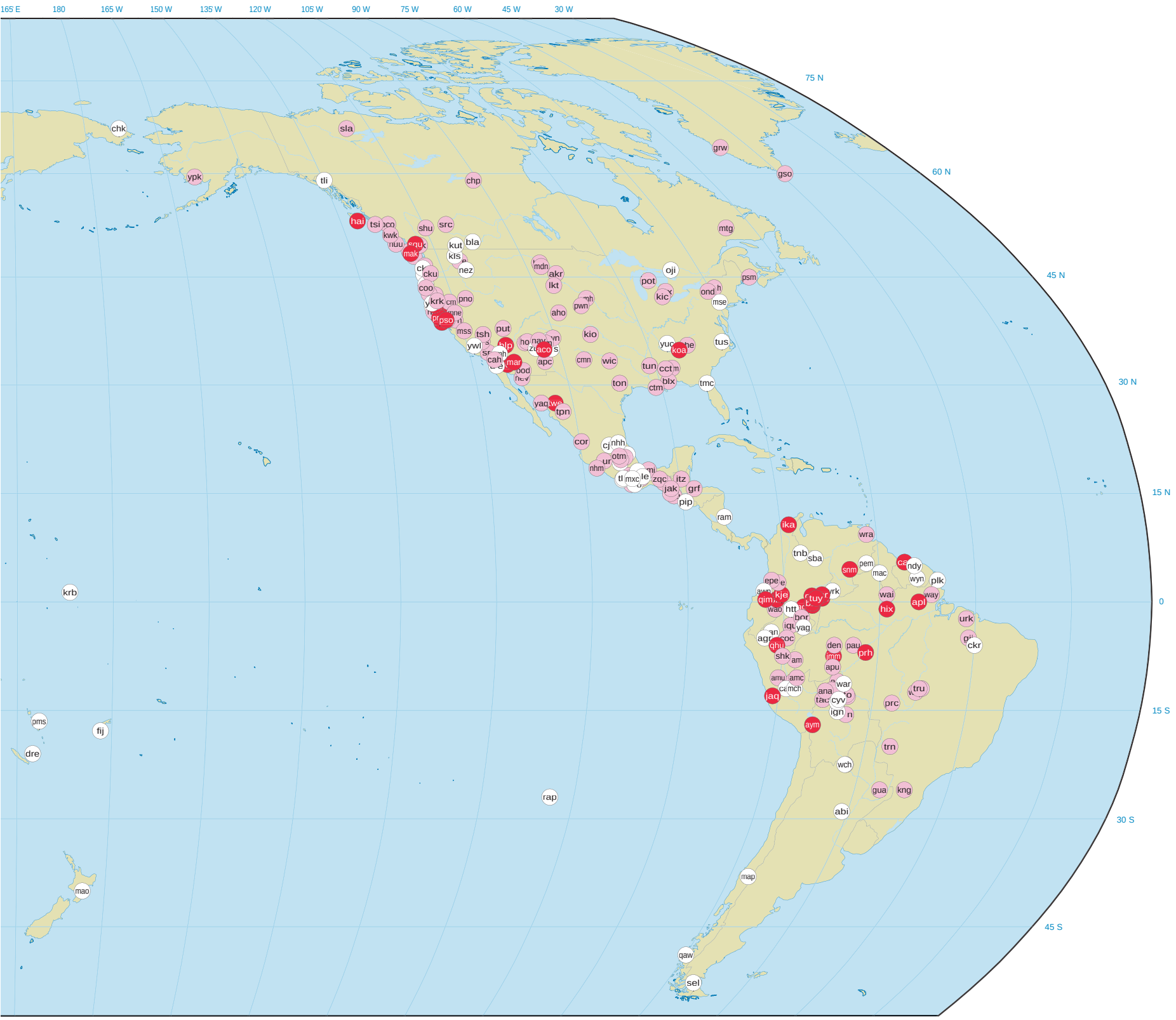


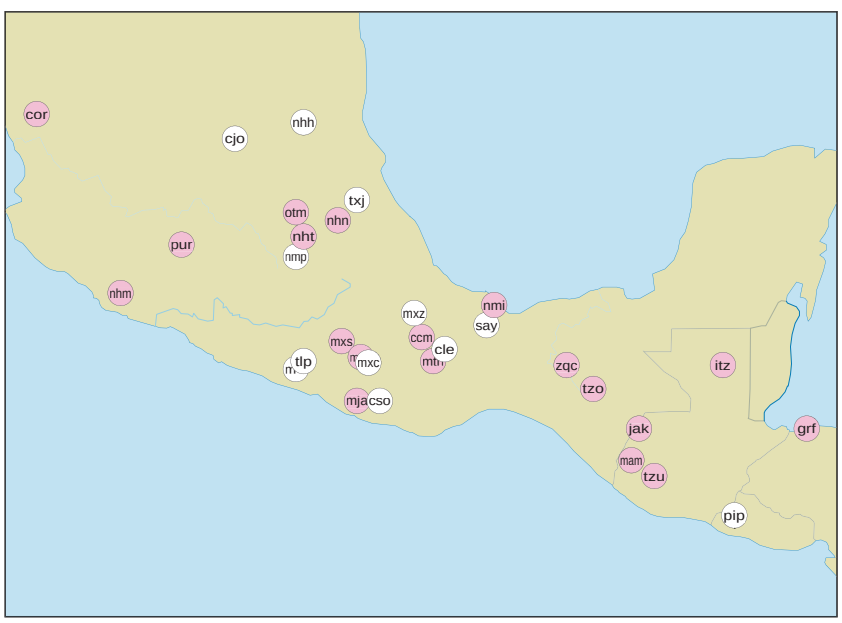


New Guinea and northern Australia

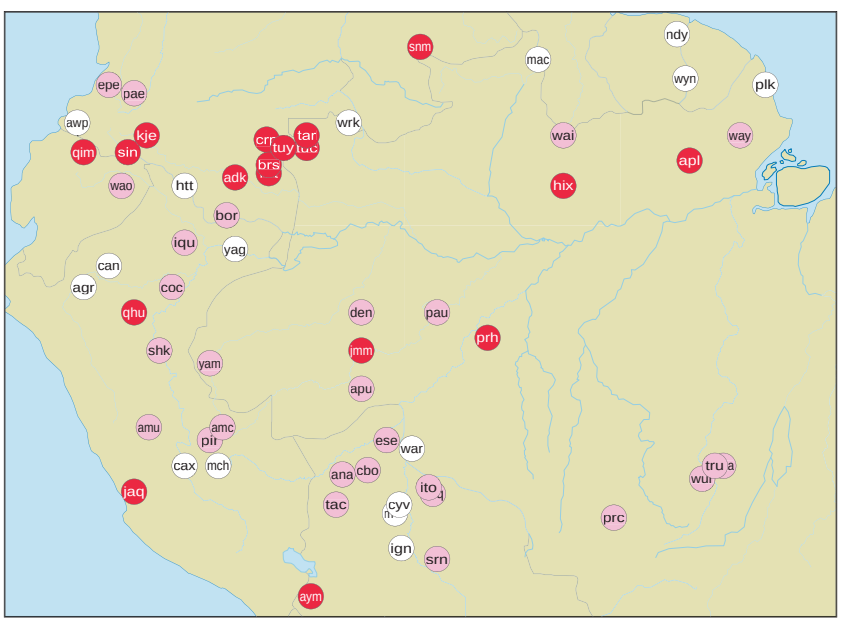


western United States





Mesoamerica



northwestern South America

78 Coding of Evidentiality

FERDINAND DE HAAN

1 Defining the values

This chapter discusses the morphological coding of evidentiality, which marks the source of information the speaker has for his or her statement. This chapter complements Chapter 77, which deals with the semantic distinctions of evidentiality. As was the case in the previous chapter, only grammaticalized evidentials are included here.

It turns out that evidentiality is marked across languages in a wide variety of ways. The following morphological means for encoding evidentiality are represented on the map:

1. No grammatical evidentials	181
2. Verbal affix or clitic	131
3. Part of the tense system	24
4. Separate particle	65
5. Modal morpheme	7
6. Mixed systems	10
total	418

These diverse means of coding evidentiality are a direct reflection of the origins of the evidentials in the respective languages. Thus, for instance, the fact that in some languages evidentiality is part of the verbal system means that these evidentials were originally tense morphemes. The same is true for the other ways of encoding evidentiality.

We turn now to a discussion of the different ways evidentiality is encoded in the sample.

From the accompanying map it appears that expressing evidentiality as **a verbal affix or clitic** is the most common strategy. With the exception of Africa it occurs on every continent. Example (1) is from Kannada (Dravidian; Sridhar 1990: 3), where the quotative morpheme *-ante* is attached to the negative verb. In Lezgian (Nakh-Daghestanian; eastern Caucasus; Haspelmath 1993: 148) the quotative morpheme *-lda* is attached to the main verb, as in example (2).

(1) Kannada
Nimma pustaka avara hattira illav-ante.
your book he.POSS near NEG-QUOT
'(It is said that) your book is not with him.'

(2) Lezgian
Qe sobranie že-da-lda.
today meeting be-FUT-QUOT
'They say that there will be a meeting today.'

In some cases the evidential morpheme is a **clitic** rather than an affix. In a number of languages the evidential can be attached to other word classes besides the verb. An example of such a language is Takelma (Takelman; Oregon; Sapir 1922: 291–2). The evidential morpheme *-ihi?*, which functions as a quotative, can be attached to any word class. This is shown in (3):

(3) Takelma
a. *naga-ihi?*
say.AOR.3SG-QUOT
'he said, it is said'
b. *ganē-ihi?*
NOW-QUOT
'now, it is said'

In a number of languages the direct-indirect evidential distinction (these terms were defined in Chapter 77) is **part of the verbal system**. An example is shown in (4) from Turkish (Aksu-Koç and Slobin 1986), where there are two past tenses that can be used for evidential distinctions.

(4) Turkish
a. *Ahmet gel-miş.*
A. come-PST.INDIR.EVD
'Ahmet must have come.'
b. *Ahmet gel-di.*
A. come-PST.DIR.EVD
'Ahmet came.'

Most languages that use the verbal system to code evidential distinctions do so only in the past tense. Some languages, such as the Caucasian languages Mingrelian, Svan, and Tsova-Tush, have evidential distinctions in the present and future as well.

While in some languages the distinction between direct and indirect evidentiality has been grammaticalized (Turkish is such a language), this is not universally the case. In Georgian (Kartvelian), past-tense indirect evidentiality has been grammaticalized as one meaning of the Perfect, but the corresponding Aorist Past has not (yet) been formalized as a marker of direct evidentiality. An example is (5) (Boeder 2000: 285–6):

(5) Georgian
a. *tovl-i mosula*
SNOW-NOM come.PERF
'It has snowed.' (indirect evidential)
b. *tovl-i movida*
SNOW-NOM come.AOR
'It has snowed.' (neutral)

When a language uses **separate particles** for evidentiality, this is very strongly correlated with coding indirect evidentiality only. Whenever a language uses separate particles, it will only use them for indirect evidentiality. An example is (6), from Dumi (Tibeto-Burman; van Driem 1991: 263), which shows the use of a quotative particle *ʔe*:

(6) Dumi
i̶m-a mʋo: dzɨ-t-i̶ *ʔe*
he-ERG what eat-NON.PRET-S23 QUOT
'What did he/they/you say he was eating?'

The only possible exceptions in the sample to the generalization that particles are only used for indirect evidentiality are Apalaí (Carib; Koehn and Koehn 1986: 119) and Lega (Bantu; Botne 1995). In Apalaí, the particle *puh(ko)* is used to denote visual evidence. The example given is shown in (7):

(7) Apalaí
moro puh t-onah-se rohke
that VIS NONFIN-finish-CMPL only
'I could tell it was all gone.'

It is not clear, however, that this is a direct evidential, since from the translation it would appear that we are dealing with visual evidence after the fact, i.e. an inferential.

In Lega (Botne 1995: 205), the particle *ámbo*, which marks indirect evidentiality, contrasts with *ampó*, which marks direct evidentiality:

(8) Lega
a. *ámbo mû-nʋ-é ko mánɔ maku*
INDIR.EVD 2PL-drink-SUBJ 16 6.this 6.beer
'[It seems that] you may drink this beer.'
b. *ampó ɛ́kurúrá mompongɛ*
DIR.EVD 3SG.PRES.pound.FV 3.rice
'She is assuredly pounding rice [I can hear it].'

The origin of the direct evidential in Lega is a proximate pronoun, which explains the evidential's status as a particle. Deictic elements frequently serve as source material for evidentials (see de Haan 2001).

There are several instances of evidentiality coded by means of **a modal morpheme**. In many instances this element is a separate modal verb, as in (9) from Dutch, where the modal verb *moeten* 'must' can encode indirect evidentiality:

(9) Dutch
Het moet een goede film zijn.
'It is said to be a good film/ It appears to be a good film.'

In some languages the irrealis or subjunctive morpheme serves as an (indirect) evidential, as is the case in a number of Australian languages. Example (10) is from Gooniyandi (McGregor 1990: 550, Bill McGregor, p. c.), where the past subjunctive morpheme *-ja* can be an indirect evidential.

(10) Gooniyandi
Ngab-ja-widda *ngamoo-nyali.*
eat-SBJV-(3PL)NOM.ACC before-REPETITION
'They were eating here not long ago (there is evidence . . .).'

Example (11) is from Mangarrayi (Merlan 1982: 150). The past irrealis morpheme has indirect evidentiality as one of its functions.

(11) Mangarrayi
ṇajiŋ-gana *ḍo?* *a-wula-ma-ṛi* *malga* *Gumja*
place-ABL shoot IRR-3PL-AUX-PST.CONT up.to G.
'They supposedly shot from Najig right up to Gumja.'

Languages which have **more than one way of encoding evidentiality** usually have a combination of a separate particle and a verbal affix. An example is Diyari (Pama-Nyungan; Austin 1981: 173), which has a quotative particle *pinti* and an affix *-ku* which marks sensory evidence (hence a direct evidential):

(12) Diyari
a. *pinti* *nawu* *wakara-yi*
QUOT 3SG.NON.F come-PRES
'They say he is coming.'
b. *ŋapa* *talara* *wakara-la* *ŋana-yi-ku*
water rain.ABS come-FUT AUX-PRES-SENS
'It looks/feels/smells like rain will come.'

Some languages, such as Georgian and Komi-Zyrian, combine a separate evidential particle with evidential marking in the verbal system.

2 Geographical distribution

The distribution of languages with and without evidentials was discussed in Chapter 77. This section focuses on the distribution of the different formal strategies for encoding evidentiality.

The distribution of some morphological markers appears to have a geographical connection. The encoding of evidentiality in the tense system is found most often in two areas often linked to areal studies, namely the Balkans and the Caucasus. The encoding of evidentiality is a prominent feature in most Turkic languages (see Johanson 2000) and also in several Caucasian families (e.g. in Kartvelian).

The evidential use of modals is mainly a western European feature. It occurs in most Germanic languages, as well as in Finnish. In these languages evidentiality is another interpretation of modal verbs. This means of encoding occurs occasionally elsewhere, usually as part of irrealis (or subjunctive) marking (as in Australian languages such as Gooniyandi).

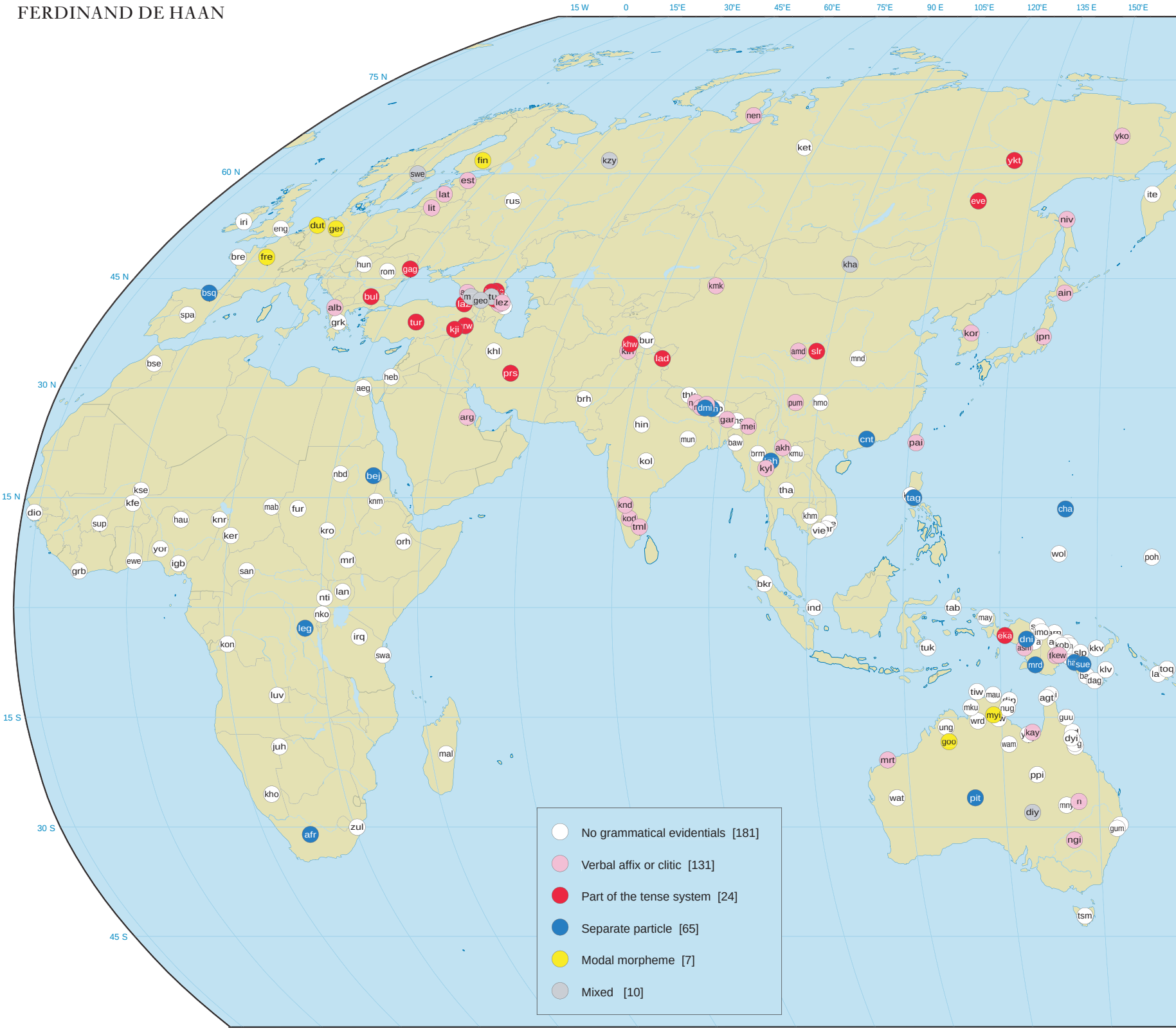
In the languages of the Americas, evidentiality is most often encoded either as a verbal affix or as a separate particle. In certain language families (e.g. Eastern Tucanoan) it is part of the tense system.

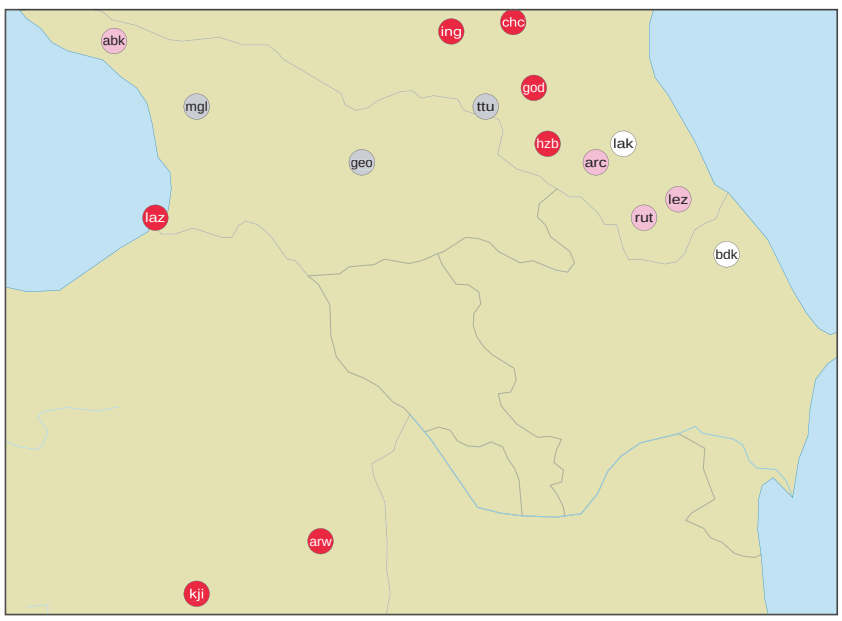
In the other areas there is little or no areal patterning discernible. It would appear that in Asia affixation on the verb is more common than any of the other means, but this is by no means a fixed rule. Whether or not areal diffusion is wholly or partially responsible is still an open question.

It has been claimed that evidentiality can be considered an areal feature (see Haarmann 1970, Aikhenvald and Dixon 1998, and Johanson and Utas 2000, among others). This claim is probably correct, given the observed clusterings of features, both semantic and morphological. From the data it seems that languages in the same geographical area can adopt structurally similar evidential notions. This means that evidentiality is a transparent category, with respect to both its semantics and its morphological coding. Evidentiality is a category that diffuses easily from one language to another, even when these languages are genetically unrelated. Of course, the fact that this can occur is no guarantee that it will occur.

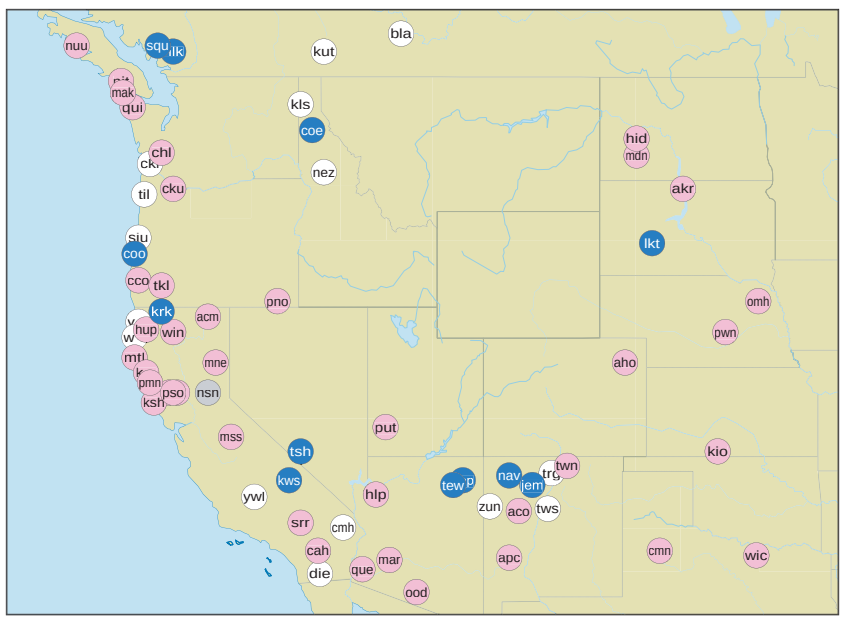
78 Coding of Evidentiality

FERDINAND DE HAAN

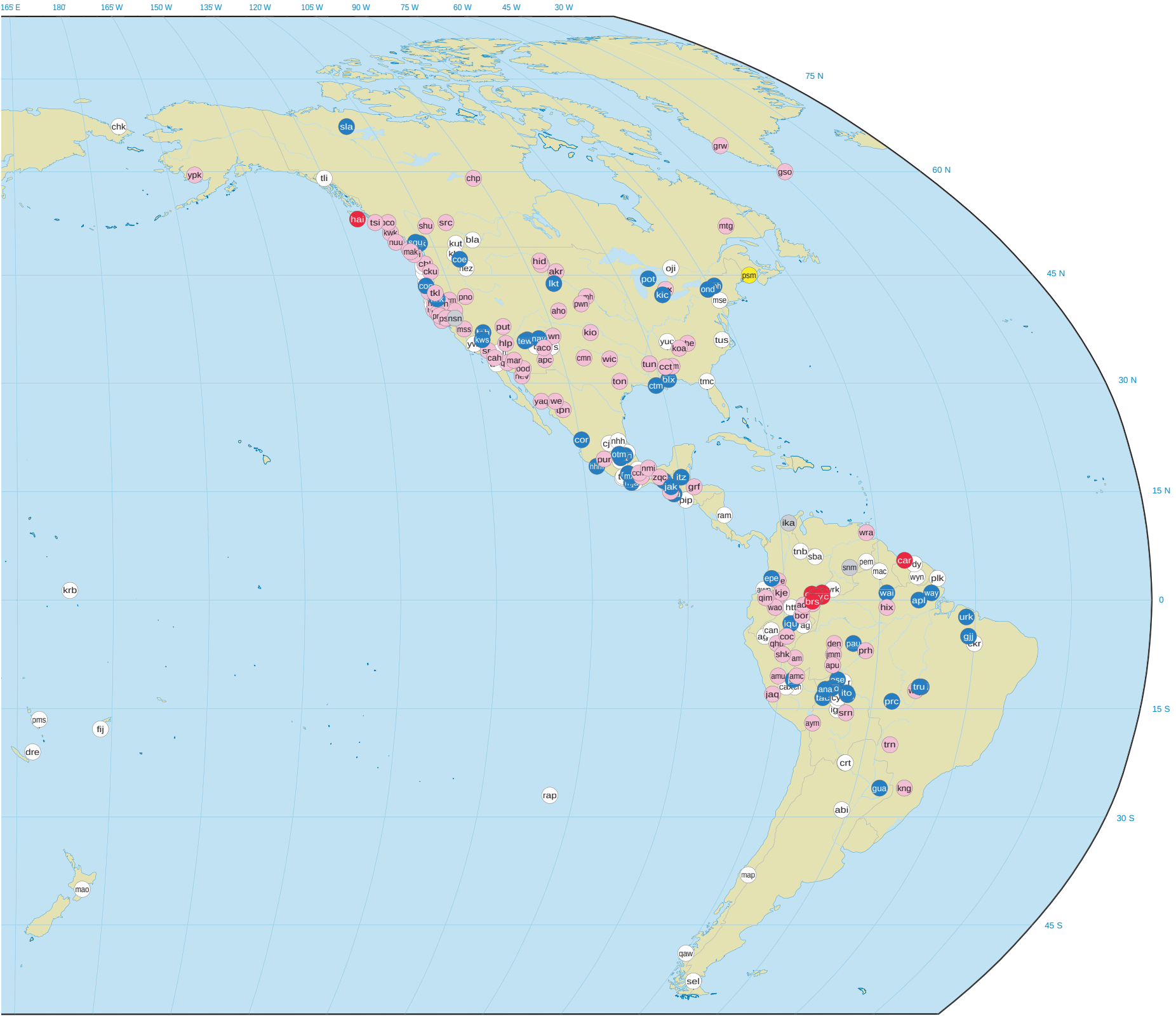


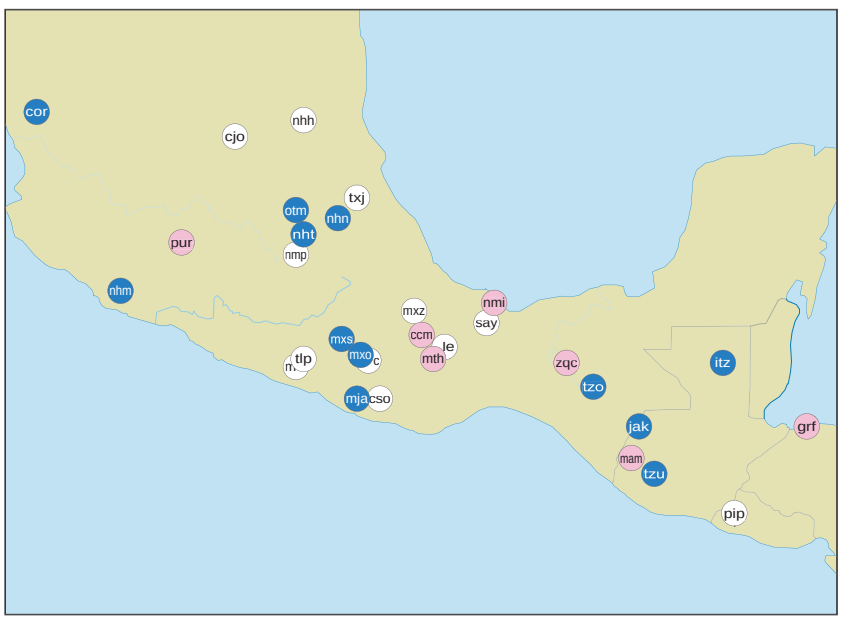


Caucasus

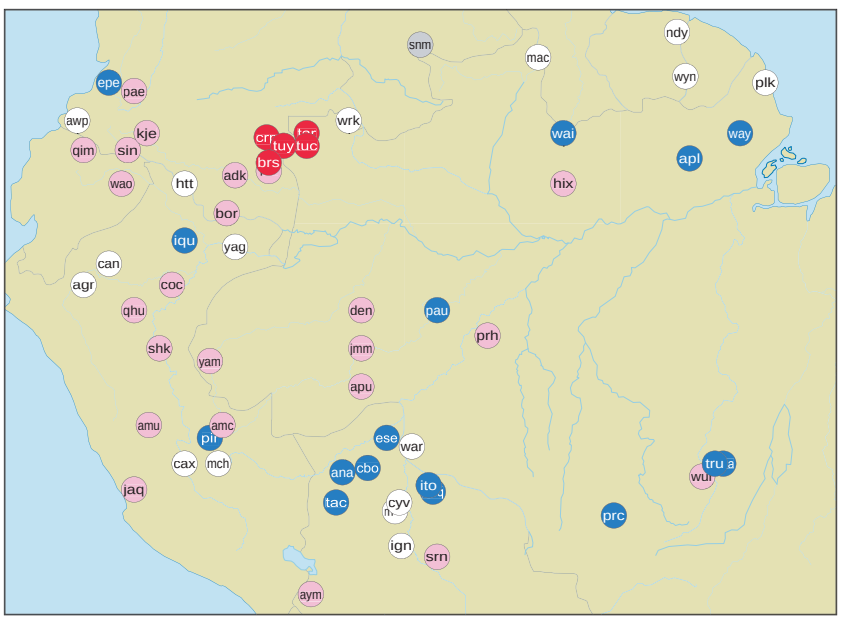


western United States





Mesoamerica



northwestern South America

79 Suppletion According to Tense and Aspect

LJUBA N. VESELINOVA

1 Suppletion in verb paradigms

The maps in this and the following chapter show the distribution of suppletion in verb paradigms according to tense-aspect, imperative and hortative moods, and verbal number. **Suppletion** is defined as the phenomenon whereby regular semantic relations are encoded by unpredictable formal patterns. Cases where the paradigmatically related forms share some phonological material are examples of **weak suppletion,** as in English *buy* versus *bought*, while cases with no shared phonological material are instances of **strong suppletion,** as in English *go* versus *went* (Dressler 1985). These types reflect two ends of a continuum rather than an either-or opposition. Mel'čuk (1994) also introduces the criterion of **uniqueness**, i.e. in order to be classified as suppletive, the alternation shown between two paradigmatically related forms should be unique, in that no other two forms in the language show exactly the same morphophonological alternation, as in Spanish *digo* 'say.PRES.1.SG' versus *dije* 'say.PRET.1.SG'. The maps concentrate on strong and unique cases.

Formal irregularity is determined by synchronic criteria. In contrast to Rudes (1980), diachronic considerations such as the historical origin of the suppletive forms and their etymologies are not taken into account here when defining the similarity or dissimilarity of paradigmatically related forms.

In the literature (Anderson 1985, Carstairs-McCarthy 1992, Carstairs-McCarthy 1994, Mel'čuk 1994) it is customary to make a distinction between **suppletion of stems,** as in English *go* 'go.PRES' versus *went* 'go.PST', and **suppletion of affixes,** as in English *boy-s* 'boy-PL' versus *ox-en* 'ox-PL', where the distribution of the plural affixes is not phonologically conditioned. The distribution of stem suppletion is presented on the maps below.

It is also customary to distinguish between **inflectional** and **derivational** suppletion. The maps exemplify both kinds. Tense and aspect as well as imperative are semantic categories associated with verbs in an inflectional way. Verbal number (or pluractionality) is a category that rarely becomes inflectional (Mithun 1988b, Corbett 2000); it is associated with verbs in a derivational and even lexical manner. Some scholars, such as Mithun, Mel'čuk (1994), and Corbett, do not accept the suppletive status of verbal pairs expressing verbal number and rather see them as separate lexical items. However, the languages where such verb pairs are found exhibit a great deal of variation as regards the obligatoriness of their use as well as their function, so that a clear-cut decision about their status in the grammar and lexicon of the languages where they occur is often difficult. No distinction between inflectional and derivational suppletion is made for these maps.

Suppletion is generally considered to be rather unsystematic and accidental, hence difficult to explain or downright uninteresting. However, several studies (Bybee 1985; Markey 1985; Deshpande 1992; Aski 1995; Fertig 1998; Veselinova 2003) have shown that this assumption is not correct. The historical development of suppletive forms provides ample evidence that they are semantically motivated (Deshpande 1992) and also well integrated in the inflectional systems of their languages (Aski 1995). The synchronic data mapped below present evidence that it is not only possible to discern patterns of suppletion according to grammatical categories, but that these patterns can be further correlated with particular groups of lexemes and show rather different areal and genealogical distributions.

2 Defining the values

The main map shows the distribution of suppletion according to tense and aspect categories. It is accompanied by an inset map where the distribution of suppletive forms in the imperative and hortative moods is presented.

The set of values on Map 79 are as follows:

1. Suppletion according to tense	36
2. Suppletion according to aspect	10
3. Suppletion in both tense and aspect	24
4. No suppletion in tense or aspect	123
total	193

English is an example of a tense-suppletion language, as in *go* versus *went*. Suppletive forms may be used to mark one tense distinction only, as in English, or to mark several tense distinctions, as in Alamblak (Sepik; Papua New Guinea).

(1) Alamblak (Bruce 1984: 146): 'go'

kit-wë-r	go-PRES-M.SG
(yi)riah-r	go.FUT-M.SG
yifi-r	go.HOD-M.SG
r-i-ë-r	HEST-go-HEST-M.SG
yi-më-r	go-REM-M.SG

The second value, which covers languages where suppletive forms are used to express mainly aspectual distinctions, is illustrated by Oneida (Iroquoian; New York State and southern Ontario).

(2) Oneida (Karin Michelson, p. c.): 'say'

kātuhe?	1.SG.say.IMPF
wakə	1.SG.say.STATIVE/PERF
wa?kīlu?	1.SG.say.PFV

Suppletive forms may be used to make the general imperfective versus perfective distinction, as in (2) above, or to express finer-grained aspectual distinctions, as in Chalcatongo Mixtec (Oto-Manguean; Oaxaca, Mexico).

(3) Chalcatongo Mixtec (Macaulay 1996: 170): 'go'

nbíí	come.HAB
bèì	come.PROG
kii	come.UNREALIZED
na-kii	PFV-come

The third group of languages, which shows suppletion in both tense and aspect, includes languages where verbs coding both of these categories are encountered, as in Lezgian (Daghestanian; eastern Caucasus). It also covers languages with more complex verbs where the stem alternation is determined by both temporal and aspectual distinctions in one and the same paradigm, as in Russian below.

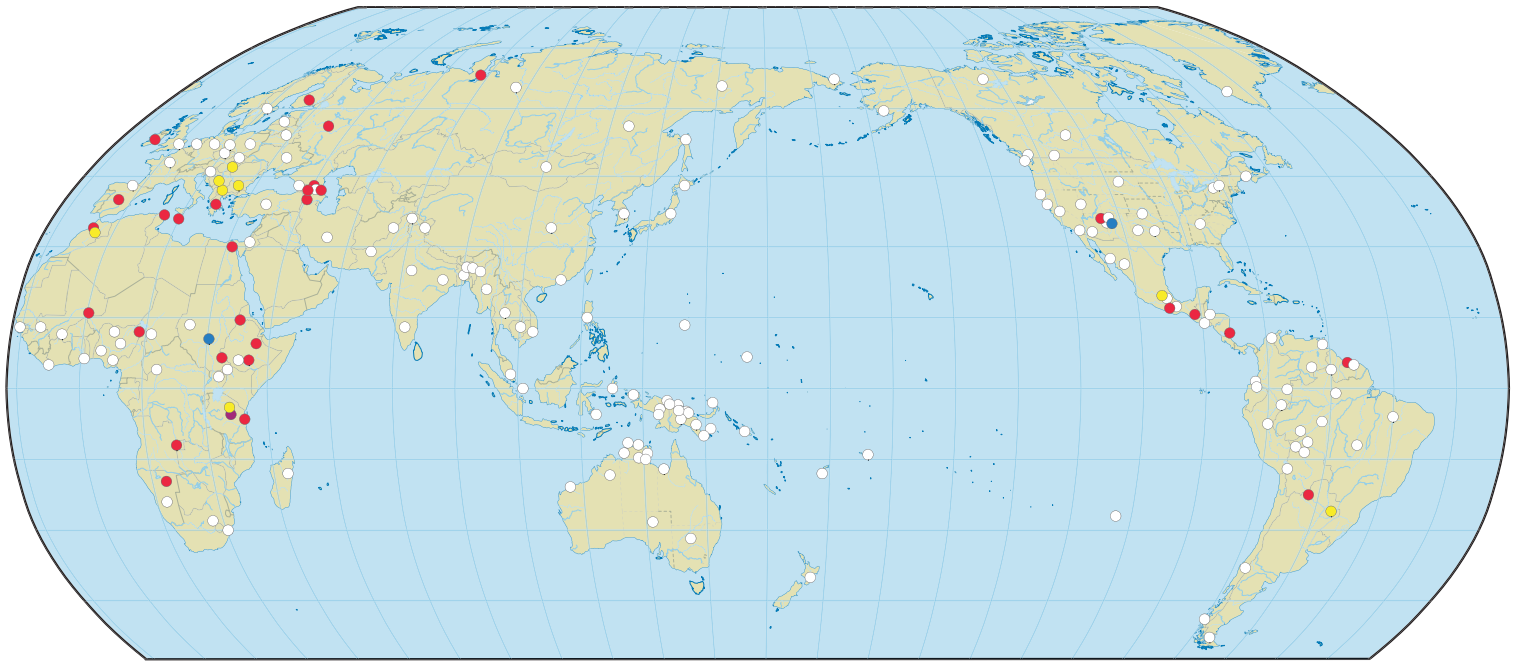
(4) Lezgian (Haspelmath 1993: 135–6): 'be' and 'do'

a. *ja*	be.PRES	b. *awú-na*	do-AOR
tir	be.PST	*ijí-zwa*	do-IMPF

(5) Russian (Maria Koptjevskaja-Tamm p. c.): 'go'

ja pri-xožu	1SG near-go.PRES-1SG
ja pri-xod-il	1SG near-go.IMPF-PST.M
ja pri-še-l	1SG near-go.PFV-PST.M
ja pri-d-u	1SG near-go.FUT.PFV-1SG
ja budu pri-xod-it'	1SG be.FUT-1SG near-go.IMPF-INF

The number of verbs per language showing tense-aspect suppletion tends to be very restricted, ranging from one single verb to ten at the most (more than three suppletive verbs are observed typically in the third group). The verb meanings encountered with all three kinds of tense-aspect suppletion are (in decreasing order of frequency) commonly used verbs such as the different kinds of copula (if it exists), 'go/come', 'say/speak', 'see', 'give/take', 'do', 'eat', and 'die'. Other verb meanings, occurring in the third group only and generally in very few languages, are 'get', 'hear', 'stay', 'throw', 'put', 'call', 'become cold', 'become', 'bear', and 'drink'.



Map 79A Suppletion in Imperatives and Hortatives

3 Suppletion in Imperatives and Hortatives (inset map)

Rudes (1980: 655) puts forth a claim that has rarely been questioned, namely, that within a given language, all suppletive verbs tend to follow the same pattern; in particular, if the language has tense-aspect-mood suppletion, this should be constant across verbs. However, the data collected here challenge this view, as is shown by the inset map on suppletive imperatives and hortatives.

1.	A regular and a suppletive form alternate	8
2.	Imperative suppletion	29
3.	Hortative suppletion	2
4.	Imperative and hortative suppletion	1
5.	No suppletion in imperatives or hortatives	153
	total	193

Values of Map 79A Suppletion in Imperatives and Horatives

The Egyptian Arabic example (6) gives an illustration of suppletion just in the imperative.

(6) Egyptian Arabic (Mitchell 1962: 97): 'come'

ʔana geet	I come.PFV.1.SG
ʔana ʔaagi	I come.IMPF.1.SG
taʕaali	come.IMP.F.SG
taʕaalu	come.IMP.PL

The suppletion here occurs only in the imperative, and only with the verb "come". This is typical. Unlike tense-aspect suppletion, suppletive forms in the imperative and hortative occur mainly with two verb meanings: 'go/come' and 'give'. Other verb in this group are those used for negative commands (a lexical verb meaning 'do not do'), as well as (but only infrequently) 'sit down' and 'be'. Suppletive imperatives and suppletive forms marking tense-aspect distinctions appear as different phenomena both synchronically (see areal distribution in §4) and diachronically. Suppletive imperatives tend to result from incorporating verbs restricted to the imperative function into the richer paradigms of semantically similar verbs, or from borrowing in situations of intensive language contact. Neither of these pathways is particularly common for the development of suppletive forms according to tense and aspect.

4 Geographical distribution

Tense-aspect suppletion shows clear areal patterns. Its mainstay is Europe and the western parts of Asia, with the South Asian subcontinent as its eastern limit. The phenomenon is competely absent in South-East Asia. Other hotbeds are found in Papua New Guinea and Mesoamerica. Tense-aspect suppletion is much more weakly represented in Africa, North and South America, and Australia.

The distribution of the different kinds of tense-aspect suppletion shows areal patterns as well. In Europe and western Asia, the dominant types are tense and tense-aspect suppletion.

The use of suppletive forms to mark tense distinctions is the dominant type of tense-aspect suppletion in Papua New Guinea.

Languages with suppletive verbs following aspectual distinctions only are markedly absent in western Eurasia. This feature appears mainly concentrated in Mesoamerica with languages from the Oto-Manguean family. Some Southern Uto-Aztecan languages in this area show verbs with this feature as well as verbs displaying both tense and aspect suppletion. The few instances of tense-aspect suppletion in other parts of the North American continent are again of the aspectual kind, occurring in Slave (Athapaskan; North-West Territories, Canada) and Oneida (Iroquoian).

In South America, the marking of tense distinctions by suppletive forms is found with all three Carib languages on the map as well as with the only Barbacoan language represented here; the use of suppletive forms to mark aspectual distinctions only is observed in Ika (Chibchan; Colombia).

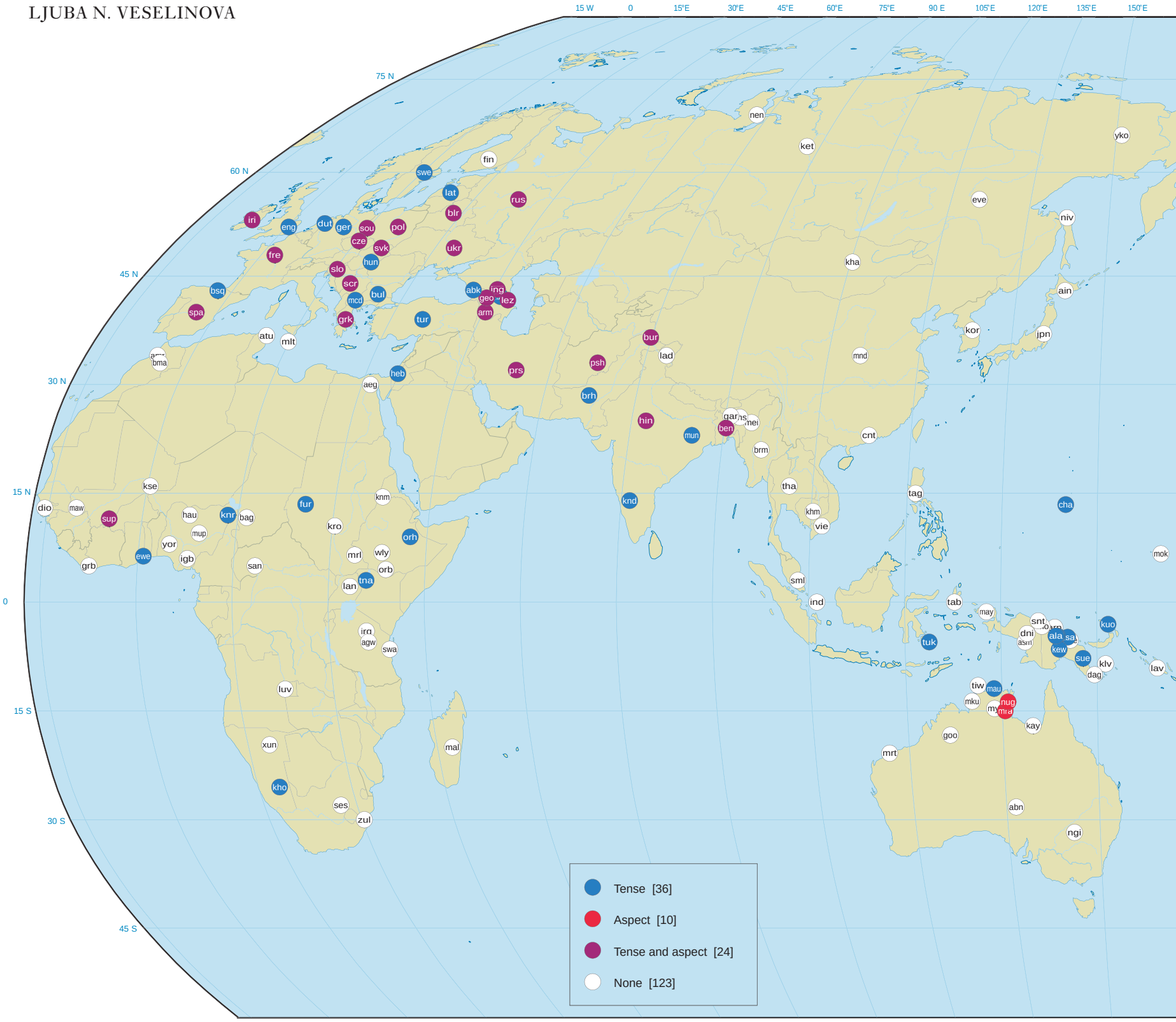
In Australia, tense-aspect suppletion is observed with three languages: Maung, Mara, and Nunggubuyu, all in the northern part of the continent, and all non-Pama-Nyungan.

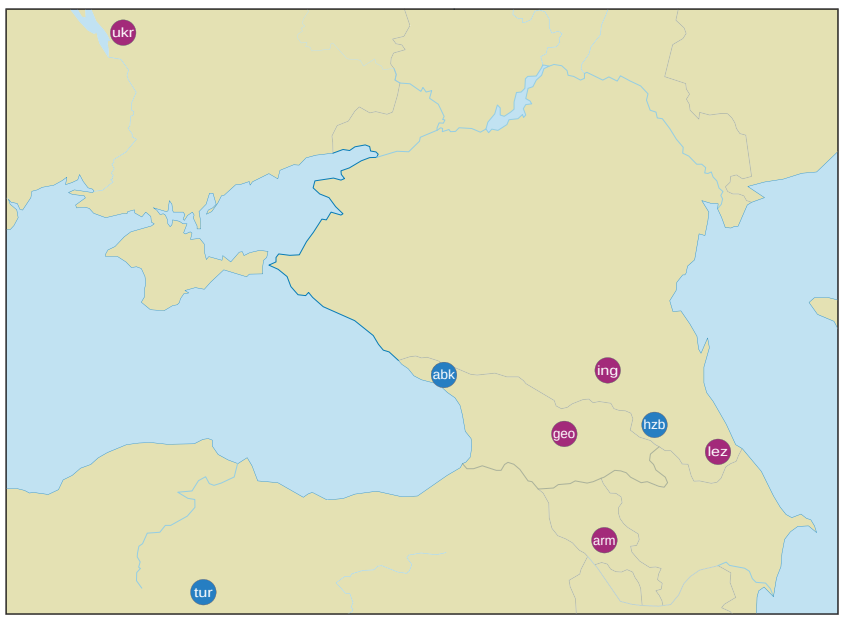
Suppletion according to tense is represented by a limited number of languages in sub-Saharan Africa. It occurs in three Nilo-Saharan languages, Turkana (Nilotic; Kenya and Uganda), Kanuri (Saharan; Lake Chad region), and Fur (Sudan), as well as in two non-Bantu languages of the Niger-Congo family, Ewe (Kwa; Togo and Ghana) and Supyire (Gur; Mali); the latter is the only sample language in Africa where aspectual suppletive pairs occur as well. The other instances of tense suppletion in Africa occur in Harar Oromo (East Cushitic; Ethiopia) and Khoekhoe (Central Khoisan; Namibia).

Unlike tense-aspect suppletion, suppletive imperatives are hardly used in Europe except for Irish, Russian, and the languages of the Caucasus and the Balkans. Family-wise, they are present in all Uralic languages mapped here, which apart from Hungarian do not show tense-aspect suppletion. The stronghold of suppletive imperatives appears to be Africa: they are observed both in Afro-Asiatic languages north of the Sahara and in languages from all major language families in sub-Saharan Africa, the eastern parts in particular. They are also found in Mesoamerica in Jakaltek (Mayan; Guatemala) and Rama (Chibchan; Nicaragua), both languages where tense-aspect suppletion is not observed.

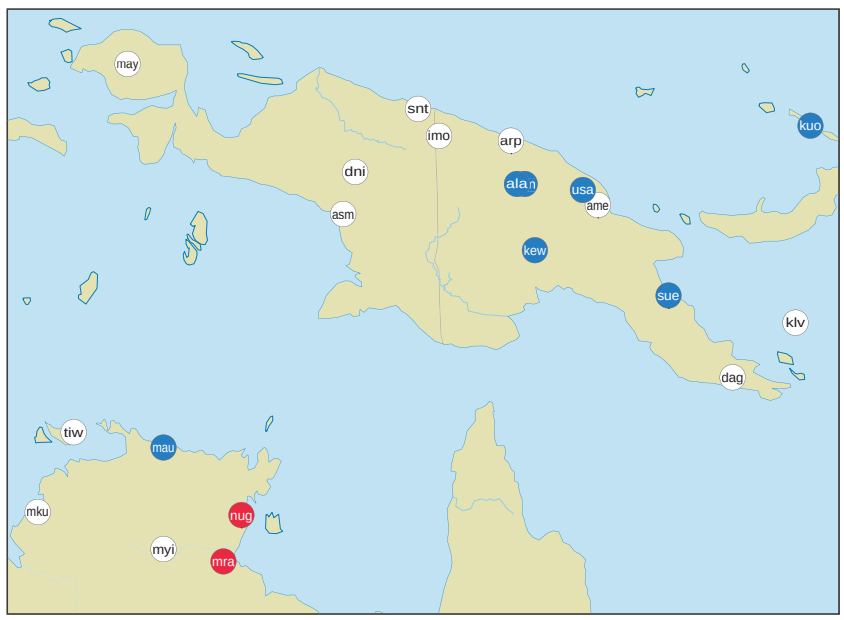
79 Suppletion According to Tense and Aspect

LJUBA N. VESELINOVA

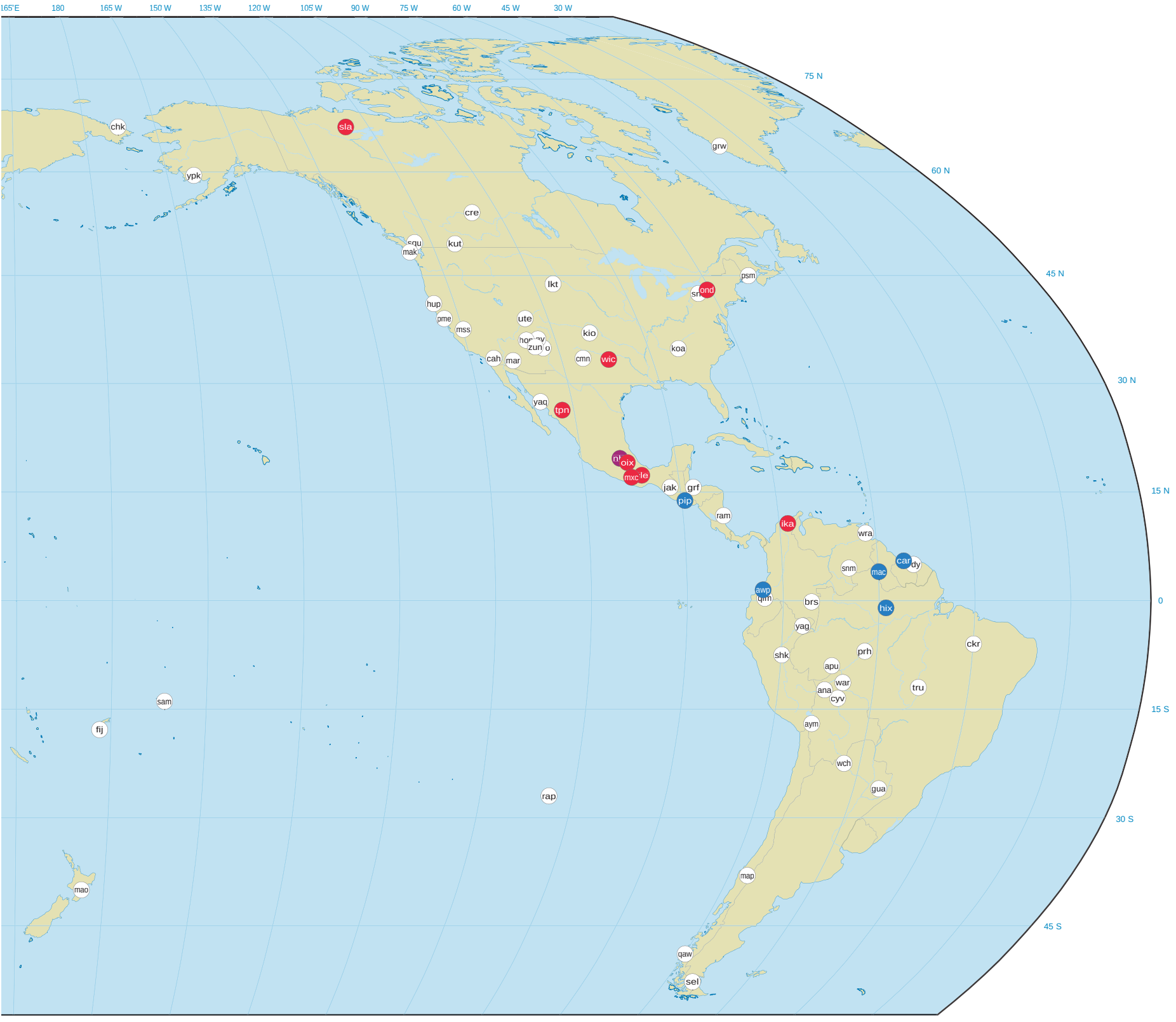


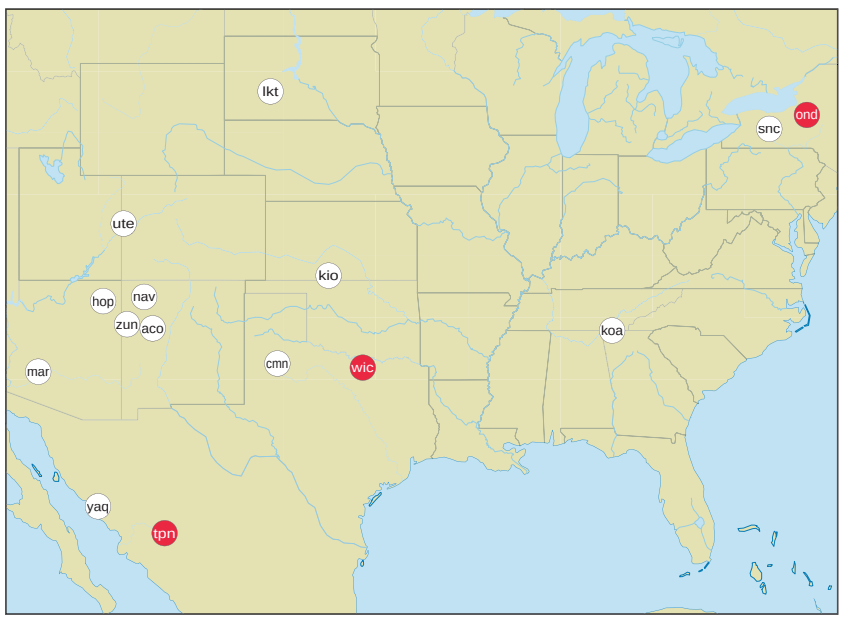


Caucasus and Black Sea area



New Guinea and northern Australia





eastern United States

Central America

80 Verbal Number and Suppletion

LJUBA N. VESELINOVA

1 Introduction

This chapter shows the distribution of suppletion according to verbal number. The term **verbal number** requires some clarification: number as a verbal category can reflect the number of times an action is performed or the number of participants in the action. When we say that a verb or verb phrase in an Indo-European language, say French, is in the **plural**, as in *ils étaient fatigués* ('they were tired'), what we mean is that the verb phrase *étaient fatigués* agrees with the subject *ils* in number. The term **plural verbs** (or **pluractional verbs**) in many of the languages to be discussed below, on the other hand, refers to any or all of the following senses: the action is performed several times (**iterative**), or at several places (**distributive**), or the action affects or involves several participants. For example, Mupun (West Chadic; Nigeria) uses derivational means such as infixation in (1a), suffixation in (1b), or completely different verbs in (1c) to express plural action or plural participants.

(1) Mupun (Frajzyngier 1993: 55–8)

a.	*pūt*	'go out.SG.ACTION'
	pú<á>t	'go.out<PL.ACTION>'
b.	*tù*	'kill.SG.ACTION'
	tù-é	'kill-PL.ACTION'
c.	*cīt*	'beat.SG.ACTION'
	nás	'beat.PL.ACTION'

The derived verbs *pú<a>t* 'go.out<PL>' and *tu-e* 'kill.PL' can be used to indicate that several people went out and that someone killed many things, respectively. However, the use of the plural verbs in (1) is not a matter of syntactic agreement. For one thing, the plural marker *mo* on the noun that would trigger agreement is itself optional, as illustrated by (2b).

(2) Mupun (Frajzyngier 1993: 60)

a.	*n-tu*	*joos*	b.	*n-tue*	*joos*	*(mo)*
	1.SG-kill.SG	rat		1.SG-kill.PL	rat	(PL)
	'I killed a rat.'			'I killed rats.'		

Secondly, while the singular form of a verb is generally ruled out with a plural object (cf. 3a), a plural verb may be used with a singular object to indicate an action performed with some intensity, or many times, as in (3b).

(3) Mupun (Frajzyngier 1993: 59)

a.	**wu*	*cit*	*mo*	b.	*wu*	*nás*	*war*
	*3.SG.M	hit.SG	3.PL		3.SG	hit.PL	3.SG.F
	*'He hit them.'				'He hit her many times.'		

Thus, while it is a separate grammatical category, verbal number, in many instances, comes very close to other categories and at times becomes indistinguishable from them. In particular, in a number of cases, verbal number is akin to aspect; there are also cases where verbal number is not easy to distinguish from nominal number as reflected in agreement (Corbett 2000: 256).

There are a few other characteristics of verbal number as a grammatical category which need to be pointed out. As indicated above, the labels *singular* and *plural* here refer to quantification of the verbal action rather than the nominal arguments of the verb. However, in some languages such as Ainu, the quantity denoted by these labels differs in yet another way from what is usually meant by 'singular' and 'plural'. In the agreement category of number, *singular* refers to 'one' and *plural* to 'more than one'. In Ainu, however, if quantity is specified by a numeral and is small, such as 'one', 'two', 'three' and sometimes 'four', the singular verbal-number form of the verb is used. For larger numbers, normally, the plural form is used. This applies both to regularly derived and to formally unrelated (but semantically related) verbal number pairs, and is illustrated by the pair 'come' in (4).

(4) Ainu (Tamura 1988: 40)

a.	*tu*	*okkaypo*	*ek*		
	two	youth	come.SG		
	'Two youths came.'				
b.	*tupesaniw*	*ka*	*arki*	*ruwe*	*ne*
	eight	even	come.PL	NMLZ	COP
	'Eight people came.'				

Finally, the expression of verbal number in languages where this category is observed may be both lexical and morphological. The main focus of Map 80 is on lexical expressions, i.e. verb pairs such as the Ainu pairs *ek* versus *arki* 'come.SG' versus 'come.PL', or the Mupun pairs *cit* versus *nás* 'hit.SG' versus 'hit.PL'. However, inasmuch as such pairs can be seen as derivationally related, the morphological expression of verbal number is taken into account as well. It should be noted that the selection of singular versus plural verb stem generally follows an ergative-absolutive pattern: for intransitive verbs, the stem reflects the number of the only argument, whereas with transitive verbs the stem reflects the number of the direct object.

(5) Ainu (Tamura 1988: 38)

a.	'stand'		b.	'kill'	
	as	'stand.SG'		*rayke*	'kill.SG.OBJ'
	roski	'stand.PL'		*ronnu*	'kill.PL.OBJ'

It is a controversial issue whether verb pairs such as the Mupun pair *cit* versus. *nás* 'hit.SG' versus 'hit.PL' are to be considered cases of suppletion according to verbal number, or simply as different words. There exists a discrepancy between grammars on the one hand, and theoretical work in morphology and suppletion on the other. While grammars usually treat pairs such as *cit* versus *nás* as cases of suppletion, authors of theoretical works (e.g. Mithun 1988b: 214, Mel'čuk 1994: 386–7, Corbett 2000: 258–9) argue that such cases are to be considered separate lexical items which are related semantically but not paradigmatically. This conflict in views seems to originate from the way one defines **suppletion**. Apparently, grammar authors are content to focus on the presence of morphological patterns of any kind (including derivational patterns), and any exceptions to these patterns are seen as suppletive. Authors of theoretical works in morphology, on the other hand, appear to restrict the term *suppletion* to cases such as *go* versus *went*, which represent clear violations of general inflectional rules. For the purposes of the current chapter, cases of suppletion are considered those which satisfy either of the following criteria: (i) exceptions to very productive derivational patterns; (ii) exceptions to established agreement patterns.

2 Defining the values

The languages of the sample were classified according to the following parameters: (i) the semantic distinction of singular-(dual)-plural made by the relevant words; (ii) whether these words can be seen as being in a relation of suppletion, in terms of the view of suppletion outlined above. The set of values on Map 80 are as follows:

1.	No singular-(dual)-plural pairs/triples in the reference material	159
2.	Singular-plural pairs, no suppletion	12
3.	Singular-plural pairs, suppletion	15
4.	Singular-dual-plural triples, no suppletion	5
5.	Singular-dual-plural triples, suppletion	2
	total	193

Mupun in (1)–(3) illustrates a language of the second type (no suppletion): the verbal number pairs do not form part of an agreement system, and the morphological means for deriving plural verbs are not only quite diverse, but also apply to a very limited number of verbs.

The third type covers languages such as Shipibo-Konibo (Panoan; Peru) and also languages such as Krongo (Kadugli; Sudan). In Shipibo-Konibo, almost all verbs have a single root and establish the singular/plural distinction only in the third person, by suffixing the plural marker *-kan*. The verbs *jo-* 'come' and *ka-* 'go' are the only two verbs that make the number distinction with all persons by using different singular/plural roots, as illustrated by examples with *jo-* 'come' in (6a–b).

(6) Shipibo-Konibo (Valenzuela 1997: 62–3)

a. *ja-Ø-ra* *Kako-nkoniax* ***jo-ke***
3-ABS-AS Caco-from.INTR **come.SG-CMPL**
'(S)he came from Caco.'

b. *ja-bo-Ø-ra* *Kako-nkoniax* ***be-kan-ke***
3-PL-ABS-AS Caco-from.INTR **come.PL-PL-CMPL**
'They came from Caco.'

Thus in Shipibo-Konibo, the verbs *jo-* 'come' and *ka-* 'go' appear as exceptions to established agreement patterns and are, therefore, described as suppletive.

In Krongo, pluractional verbs (usually expressing iterative and habitual meaning) are derived by means of prefixes or tone changes or reduplication, as illustrated in (7a–c). These processes apply to a great number of verbs. However, there are fifteen verbs which do not use any of these morphological means but rather express iterative and habitual meaning by completely unrelated stems; one of them is shown in (7d).

(7) Krongo (Reh 1985: 201–9, Mechthild Reh, p. c.)

a. prefix
SG *à-mà* 'VERB MARKER-answer'
PL *t-èe-mà* 'VERB MARKER-PL-answer'

b. tone
SG *ò-kídò-ònò* 'VERB MARKER -cut.off-SUFF'
PL *ò-kìdò-onò* 'VERB MARKER -cut.off.PL-SUFF'

c. reduplication
SG *t-áa-ʧà* 'VERB MARKER-PUNCT-go.away'
PL *à-ʧàa-ʧá* 'VERB MARKER-RDP-go.away'

d. suppletion
SG *à-pá-ánà* 'VERB MARKER-hit.SG-SUFF'
PL *à-rwà-ànà* 'VERB MARKER-hit.PL-SUFF'

In Krongo the functions of verbal number are very close to those of verbal aspect, such as iterative (i.e. repeated action) and habitual; the singular counterparts of the derived verbs usually express punctual action. The derivational devices illustrated above are, according to the grammatical description, lexically conditioned. However, the derivation as a whole is a very widespread one and the suppletive verbs illustrated in (7d) are easily perceived as exceptions to a pattern and hence as paradigmatically related, rather than simply as different words.

The fourth type is illustrated by Navajo (Athapaskan; Arizona and New Mexico). In this language, we observe verb triples of the kind SG.DU.PL as well as verb pairs of the kind SG.PL.

(8) Navajo (Young and Morgan 1987: 143, 133)

a. 'go, walk'
-yá- 'go.SG'
-'áázh- 'go.DU'
-áát- 'go.PL'

b. 'die'
-tsą́- 'die.SG'
-ná- 'die.PL'

In languages where verb triples of the kind SG.DU.PL are observed, verb pairs of the kind SG.PL are also found; languages with triples only do not seem to exist. As to whether these verbs are in a relation of suppletion or not, Navajo represents languages where such verbs select, rather than agree with, their arguments (Durie 1986: 358), as illustrated by (9).

(9) Navajo (Durie 1986: 358)

a. *shí* *ashkii* *bi-ł* *yi-sh-'ash*
I boy him-with PROG-1SG-walk.DU
'I am walking with the boy.'

b. *nihí* *ła'* *di-iid-áát́*
we subset FUT-1NON.SG-walk.SG
'One of us will go.'

The agreement affixes *-sh-* '1.SG' in (9a) and *-iid-* '1 NON-SG' in (9b) follow the number of the subject, whereas the selection of stem follows the number of participants. Thus, postulating suppletion for the alternation of stems such as *-'ash* and *-áát́* is not justified.

The fifth and final type is illustrated in (10) with verbs from Slave (Athapaskan; North-West Territories, Canada).

(10) Slave (Rice 1989: 790–2)

a. 'be seated'
wʰida 'be seated.SG'
wʰíkē 'be seated.DU'
c'déhkʷ'i 'be seated.PL'

b. 'die'
łaniⁿwe 'die.SG'
łakeniⁿdé 'die.DU/PL'

Although closely related to Navajo, Slave behaves differently with respect to agreement and stem selection. The alternation of stems not only reflects the number of participants, but also agrees in number with the agreement affixes.

(11) Slave (Rice 1989: 271, 281; Keren Rice, p. c.)

a. *thí-ke*
ASPECT/1DU-sit.DU
'We (DU) are sitting.'

b. *de-í-kʷ'i*
ASPECT-1PL-sit.PL
'We (PL) sat.'

As a general matter, it is important in Slave that the verb stem should agree in number with the overt subject marker. Combining a nonsingular subject marker with a singular stem, as in the Navajo example (9), is impossible in Slave for first- and second-person subjects (Keren Rice, p. c.). Since there is an established agreement pattern, the verb triples/pairs in Slave can legitimately be described as paradigmatically related and thus suppletive, given the view of suppletion adopted in this chapter.

The number of verbal number pairs/triples per language varies from one to as many as eighteen, as in Ekoka !Xun (Northern Khoisan; Namibia); however, in more than half of the languages where they occur, they range between one and four. Semantically, such verbs tend to centre around the basic motion verbs 'go/come', 'walk', and 'fall', position verbs such as 'lie', 'sit', and 'stand', as well as their transitive counterparts 'lay' and 'put'; very frequent are 'die' and 'kill', less common 'sleep'.

3 Geographical distribution

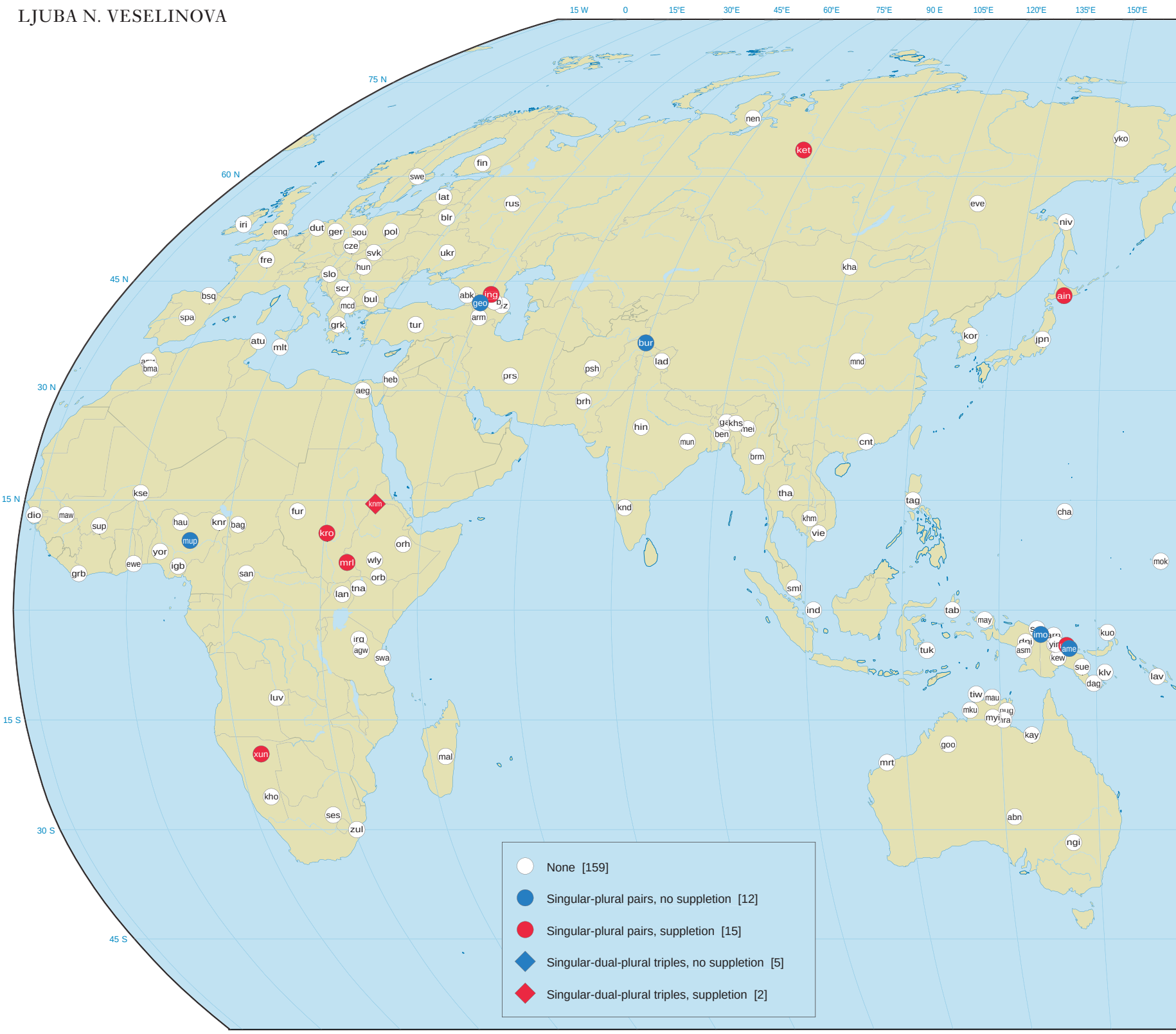
Verbal number pairs/triples are very common in North America. In South America, they occur in languages from the Macro-Ge, Panoan, and Tucanoan families. In the Pacific region, verbal number pairs are found in Samoan (Polynesian) as well as in many languages of the Trans-New Guinea family, specifically the Northern branch and languages from the Madang-Adelbert Range branch. The phenomenon is represented in all four major families of Africa, but most strongly in the Nilo-Saharan language family, specifically in subgroups such as Eastern Sudanic and Kunama (both cover languages spoken in Sudan and Ethiopia). In Asia, verbal number pairs/triples are observed in Ainu as well as Ket and Burushaski. The western limit of this phenomenon appears to be the Caucasus.

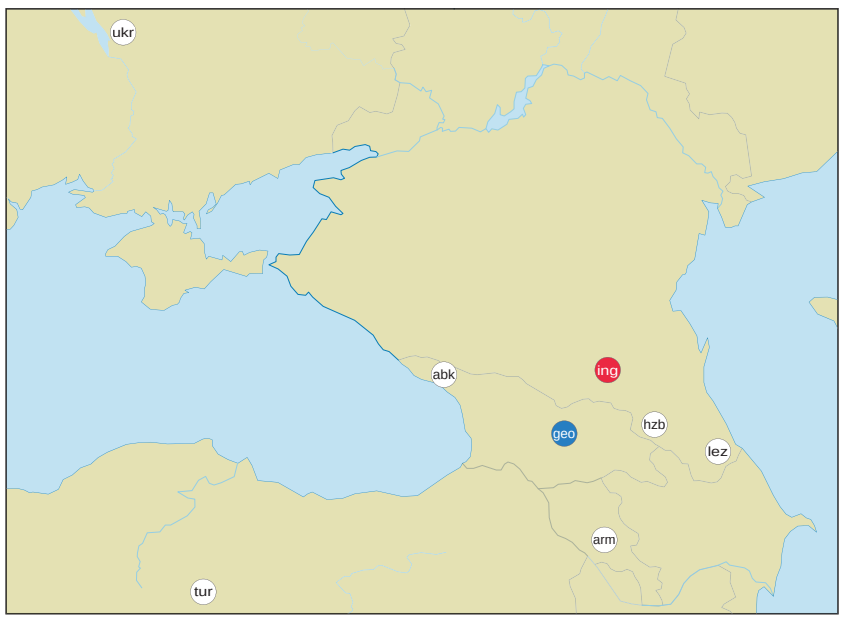
4 Theoretical issues

This phenomenon bears on several matters of theoretical and descriptive importance: the category of verbal number and its various expressions, as treated in Gil (1993) and Corbett (2000) as well as Mithun (1988b) from a historical perspective; the distinction between semantic selection and syntactic agreement, discussed in Durie (1986); the notion of suppletion (Mel'čuk 1994); and the emergence of paradigmatic relations between semantically related words (Veselinova 2003).

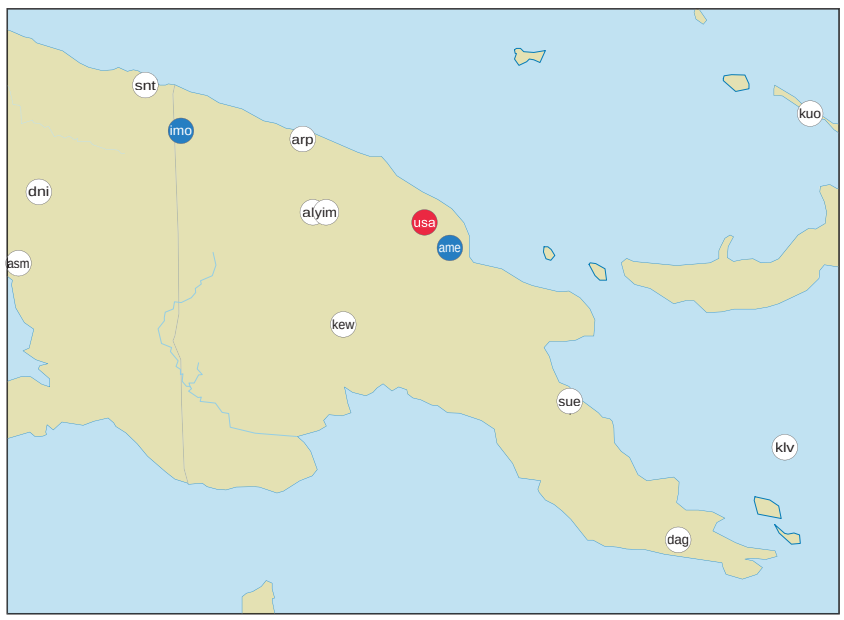
80 Verbal Number and Suppletion

LJUBA N. VESELINOVA

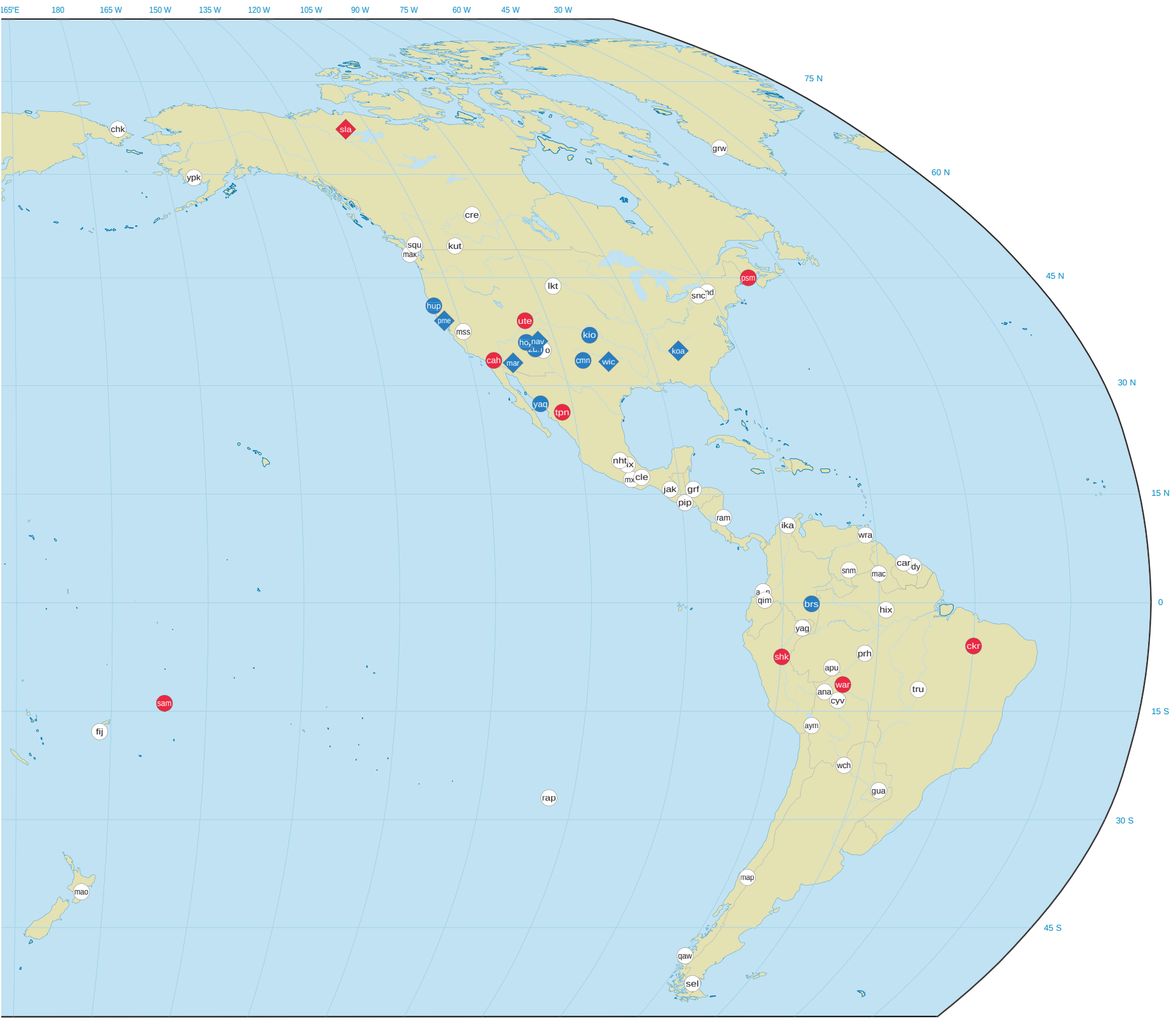


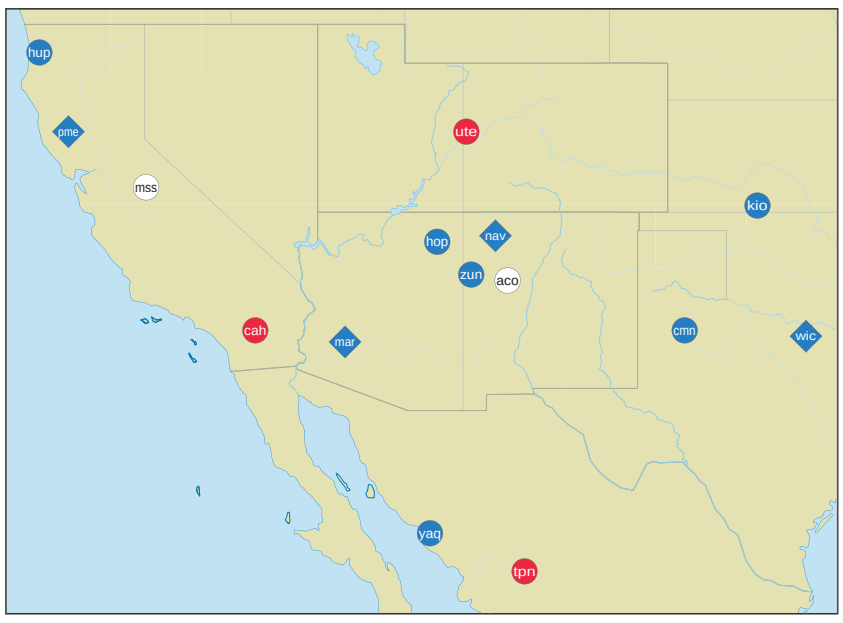


Caucasus and Black Sea area

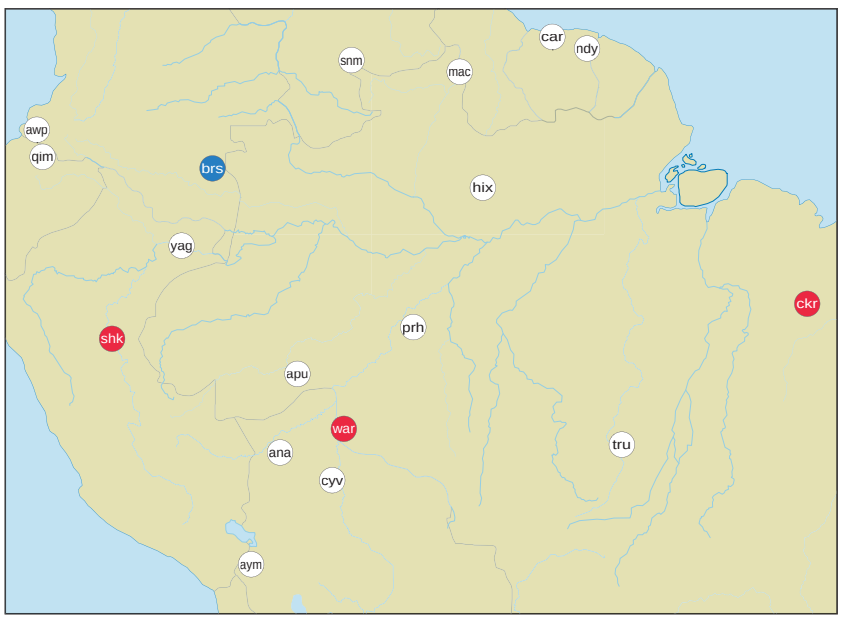


New Guinea





western North America



northwestern South America

81 Order of Subject, Object, and Verb

MATTHEW S. DRYER

1 Defining the values

This map shows the ordering of subject, object, and verb in a transitive clause, more specifically declarative clauses in which both the subject and object involve a noun (and not just a pronoun), as in the English sentence in (1).

(1) *[The dog] chased [the cat].*
S V O

English is SVO (Subject–Verb–Object), because the subject *the dog* in (1) precedes the verb while the object *the cat* follows the verb.

There are six logically possible orders of the three elements S, O, and V, as shown in the feature-value box.

1. Subject-Object-Verb (SOV)	497
2. Subject-Verb-Object (SVO)	435
3. Verb-Subject-Object (VSO)	85
4. Verb-Object-Subject (VOS)	26
5. Object-Verb-Subject (OVS)	9
6. Object-Subject-Verb (OSV)	4
7. Lacking a dominant word order	172
total	1228

All six of these types are attested; examples of each type are given in (2).

(2) a. Japanese (Kuno 1973: 10)
John ga tegami o yon-da.
John SUBJ letter OBJ read-PST
S O V
'John read the letter.'

b. Mandarin (Li and Thompson 1981: 217)
Zhāngsān shōudǎo-le yi-fēng xìn.
Zhangsan receive-PERF one-CLF letter
S V O
'Zhangsan received a letter.'

c. Irish (Dillon and Ó Cróinín 1961: 166)
Léann [na sagairt] [na leabhair].
read.PRES the.PL priest.PL the.PL book.PL
V S O
'The priests are reading the books.'

d. Nias (Austronesian; Sumatra, Indonesia; Brown 2001: 538)
i-rino vakhe ina-gu
3SG.REALIS-cook ABS.rice mother-1SG.POSS
V O S
'My mother cooked rice.'

e. Hixkaryana (Carib; Brazil; Derbyshire 1979: 87)
toto y-ahosɨ-ye kamara
man 3:3-grab-DISTANT.PST jaguar
O V S
'The jaguar grabbed the man.'

f. Nadëb (Vaupés-Japurá; Brazil; Weir 1994: 309)
awad kalapéé hapɨ̈h
jaguar child see.IND
O S V
'The child sees the jaguar.'

Although all six of these orders are attested, the last two types, OVS and OSV, in which the object comes first, are rare.

The terms *subject* and *object* are used here in a rather informal semantic sense, to denote the more agent-like and more patient-like elements respectively. Their use here can be defined in terms of the notions S, A, and P, where the S is the single argument in an intransitive clause, the A is the more agent-like argument in a transitive clause, and the P is the more patient-like argument in a transitive clause. For the purposes of this map, then, the term *subject* is used for the A while the term *object* is used for the P. A language shown on the map as SOV could thus also be equally well and perhaps more accurately described as APV. Note that many linguists use the terms *subject* and *object* somewhat differently from this, and some linguists question the applicability of these terms to some languages, but these issues do not arise with the use of these terms here. For example, there is controversy surrounding the question of what ought to be considered the subject in Philippine languages, like Cebuano (cf. Schachter 1976). Cebuano has two common ways to express transitive clauses, one of which is illustrated in (3).

(3) Cebuano (Austronesian; own data)
gi-palit [sa babayi] [ang saging]
GOAL.FOC-buy NONTOP woman TOP banana
'The woman bought the bananas.'

While there is a question as to which of the two arguments in (3) should be considered a subject (or whether neither or both should), in both types of clauses the verb normally comes first, followed by the A, and then the P. Hence, by the use of *subject* and *object* assumed for this map, Cebuano is treated as a VSO language.

Note that while the position of the subject in intransitive clauses is generally the same as in transitive clauses, in some languages this is not the case. See Chapter 82.

Some languages can be assigned straightforwardly to one of the six types, because all orders other than one are either ungrammatical or used relatively infrequently and only in special pragmatic contexts. Such languages can be said to have **rigid order**. There are many other languages in which all six orders are grammatical. Such languages can be said to have **flexible order**. Flexible order languages are sometimes described as having "free" word order, though this is misleading, since there are often pragmatic factors governing the choice of word order.

We can further distinguish two subtypes of languages with flexible word order. In some languages with flexible order, there is one order which is most common and which can be described as the **dominant order**. In other flexible order languages, the flexibility is greater and there is no one order that is the dominant order in terms of frequency of usage or pragmatic neutrality. Flexible order languages in which one order is dominant are shown on the map according to that dominant order – in other words, the map does not distinguish rigid order languages from flexible order languages with a dominant order. Flexible order languages lacking a dominant order are shown on the map as "lacking a dominant word order". Russian is an example of a language with flexible word order in which SVO order can be considered dominant, so Russian is shown on the map as SVO. See the box section "Determining Dominant Word Order" on p. 371.

There are a number of different subtypes of languages lacking a dominant order which are not distinguished on the map. In some languages with highly flexible word order, all or most orders of subject, object, and verb will be possible and common. Nunggubuyu (Gunwinyguan; northern Australia) is an example of such a language (Heath 1984: 507–13; 1986). But some languages lack a dominant order only because just the subject or just the object exhibits flexibility with respect to the verb. For example, Syrian Arabic allows both SVO and VSO orders and there does not seem to be a reason (at least on the basis of the description by Cowell 1964: 407, 411) to consider one of them dominant. However, only these two orders are common and the order of verb and object is relatively inflexible.

A third subtype of language lacking a dominant order consists of languages in which different word orders occur but the choice is

syntactically determined. For example, in German and Dutch, the dominant order is SVO in main clauses lacking an auxiliary and SOV in subordinate clauses and clauses containing an auxiliary (see Chapter 83 for examples). Because this results in both orders being common, neither order is considered dominant here and these two languages are shown on the map as lacking a dominant word order. In general, if the word order varies according to whether there is an auxiliary verb, the language is shown on the map as lacking a dominant order. Another language whose word order depends both on whether there is an auxiliary and whether the clause is a main clause is Dinka (Nilotic; Sudan): like German, the order is SVO in main clauses without an auxiliary, SAuxOV in main clauses with an auxiliary, but it is VSO in subordinate clauses without an auxiliary and AuxSOV in subordinate clauses with an auxiliary (Nebel 1948: 9, 25, 42, 75, 82).

Where languages differ in their order between main clauses and subordinate clauses, the order in main clauses is used to classify them on this map. For example, Quileute (Chimakuan; Washington State) is VSO in main clauses and SVO in subordinate clauses (Andrade 1933: 278), and is shown on the map as VSO. In some languages, word order is more fixed in subordinate clauses. For example, in Miya (Chadic; Nigeria), while both SVO and VOS are found in main clauses, only VOS order is found in adverbial subordinate clauses and relative clauses (Schuh 1998: 281, 291); because both SVO and VOS are common in main clauses, Miya is shown on the map as lacking a dominant order.

2 Geographical distribution

The most frequent of the six orders is SOV and it is widely distributed across the globe. Perhaps the most striking region in which SOV predominates is an area covering most of Asia, except in South-East Asia and the Middle East. It is also overwhelmingly the dominant order in New Guinea, most of the exceptions being along the north coast. It is the most common order among languages in Australia which have a dominant order at all, although even in languages in which SOV is dominant, the order is generally flexible. It is clearly the dominant order in North America outside of the Pacific North-West and Mesoamerica.

The map shows three areas where SVO order predominates: (i) an area covering much of sub-Saharan Africa, though with a scattering of SOV and VSO languages; (ii) an area extending from China and South-East Asia south into the Austronesian languages of Indonesia and the western Pacific; and (iii) Europe and around the Mediterranean. SVO order is not common outside these areas.

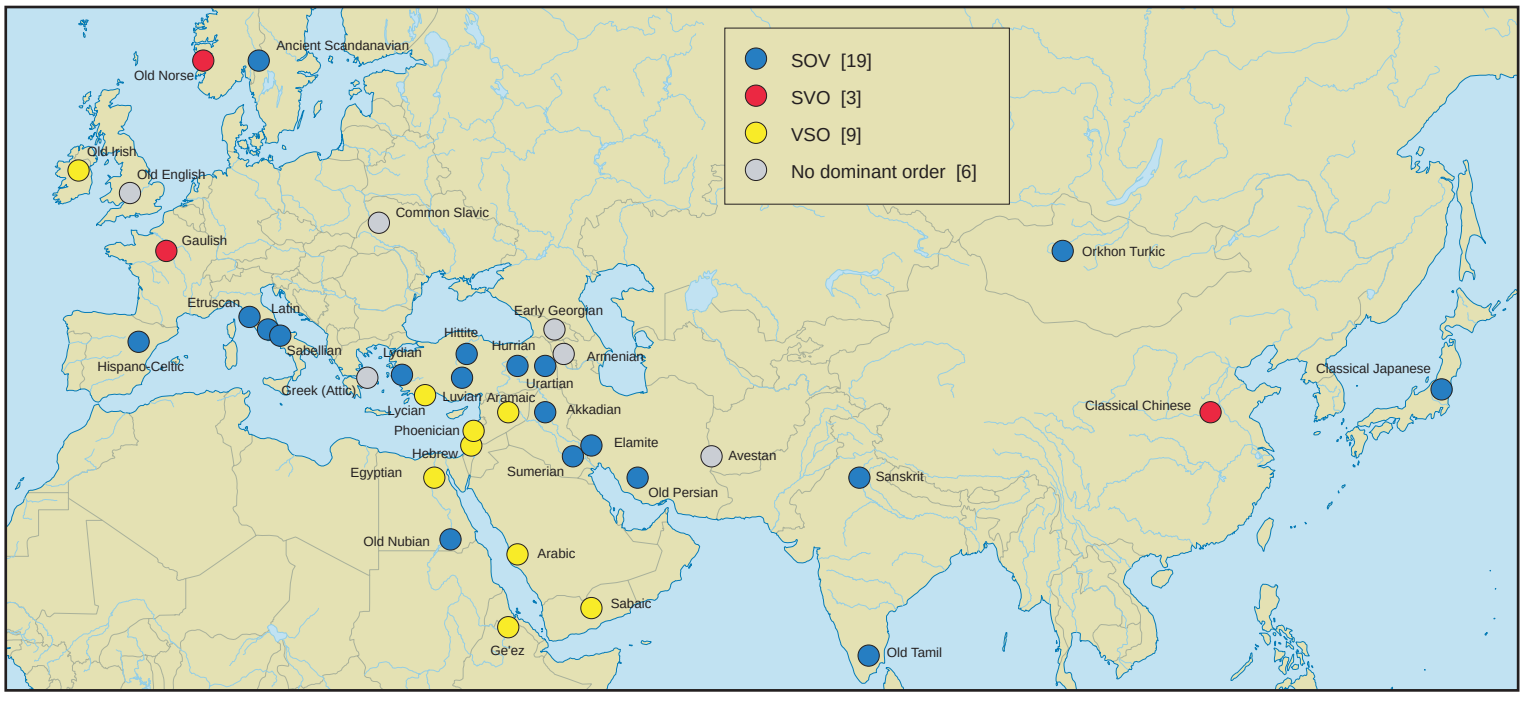
VSO order is scattered around various parts of the world, in eastern Africa (among various Eastern Sudanic languages), in North Africa (Berber), in the western extremes of Europe (Celtic), in and around the Philippines, among Polynesian languages of the Pacific, in Mesoamerica, and in the Pacific North-West. VOS order is also scattered around the globe, though there are no attested instances on the mainland of Africa or Eurasia.

There are nine OVS languages on the map, six of which are spoken in South America, five in the Amazon basin, and one (Selknam) in Tierra del Fuego. There are only four OSV languages shown: Warao in Venezuela, Nadëb in Brazil, Wik Ngathana in northeastern Australia, and Tobati in West Papua, Indonesia. Languages without a dominant order are especially common in North America and Australia, and to a lesser extent in South America. The scattering of this type partly reflects the fact that this is not a homogeneous type, since it mixes languages with highly flexible order with languages which have more rigid order but where there are two dominant orders. The former type, languages with highly flexible order, is most common in North America and Australia and relatively uncommon in Africa, Europe, Asia, New Guinea, and among Austronesian languages.

Map 81A shows the distribution of word order in various languages of the past, though the times at which these languages were spoken vary from 4500 years ago to 1000 years ago. The map illustrates the fact that SVO, now a common order in Europe and around the Mediterranean, was less common in the past: on the one hand, there were SOV languages like Latin and Etruscan in western Europe; on the other hand, there were many VSO languages in what is now the Middle East, represented both by Semitic languages and by Egyptian.

3 Theoretical discussion

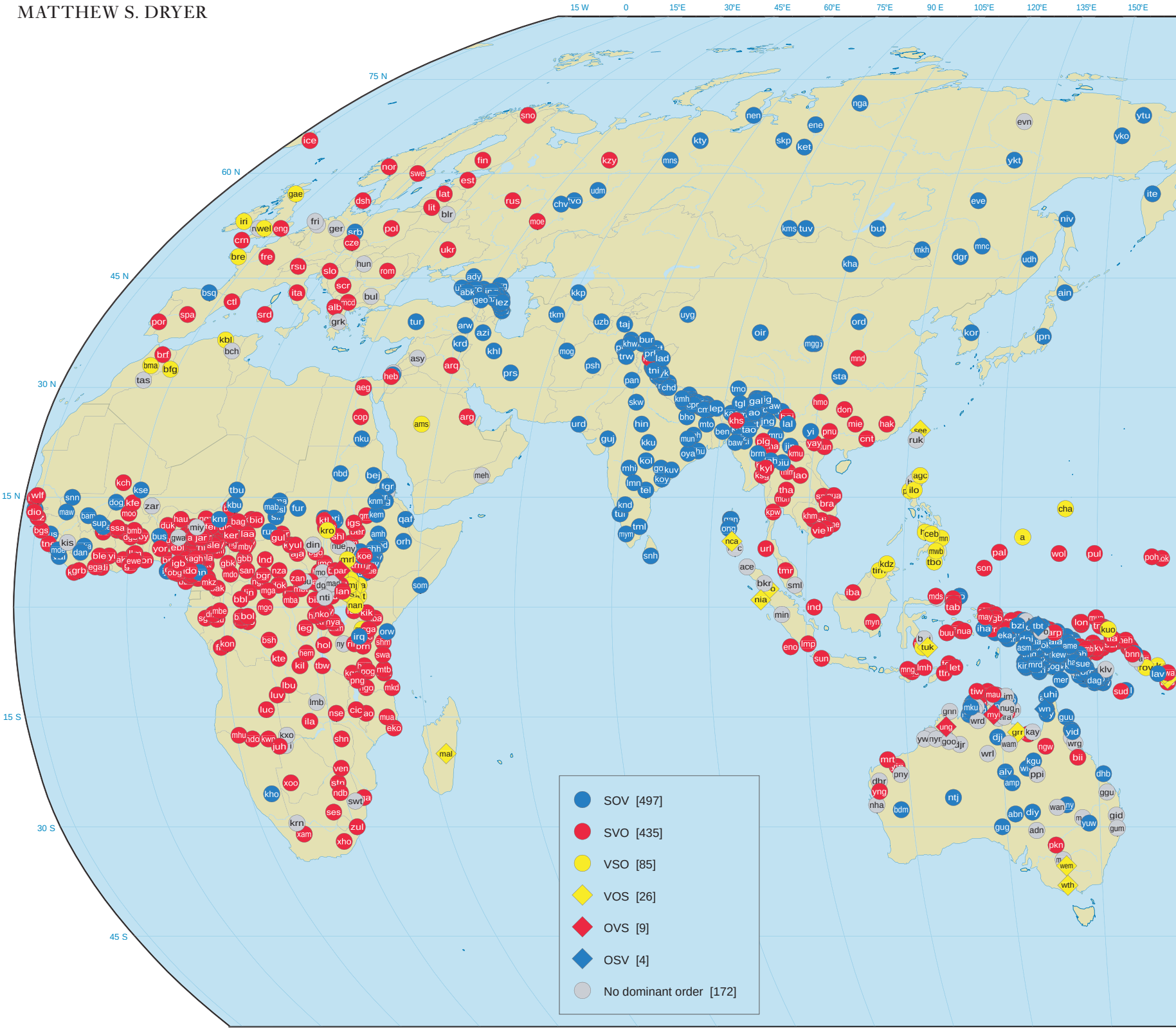
While the feature shown on Map 81 is perhaps the single most frequently cited typological feature of languages, it is now recognized that it represents a clause type that does not occur especially frequently in spoken language; it is more common that at least one of the two arguments of a transitive clause will be pronominal, and in many languages pronominal subjects are expressed by verbal affixes. It is argued by Dryer (1997) that a more useful typology is one based on two more basic features, whether the language is OV or VO and whether it is SV or VS; these are shown on the next two maps, in Chapters 82 and 83. In addition, as noted above, the order in transitive clauses is not always the same as the order in intransitive clauses. The feature shown on this map is also important in that many other features are predictable from it, at least statistically. Most of these features correlate more specifically with the order of object and verb (Greenberg 1963, Dryer 1992; see Chapters 95–97). For a few features, SVO languages exhibit properties intermediate between those of SOV languages and those of verb-initial languages, though in general they are more similar to verb-initial languages (Dryer 1991).

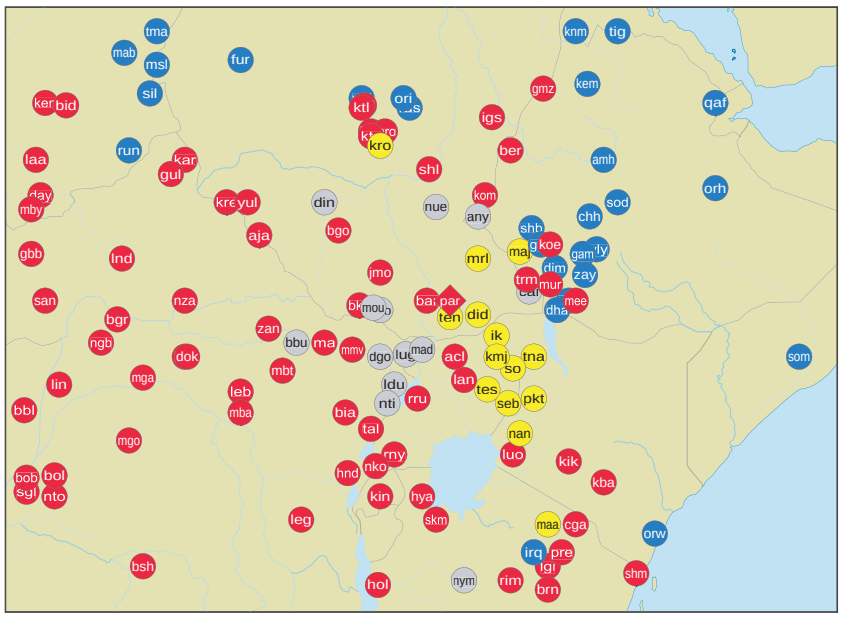


Map 81A Order of Subject, Object, and Verb in "Ancient" Languages

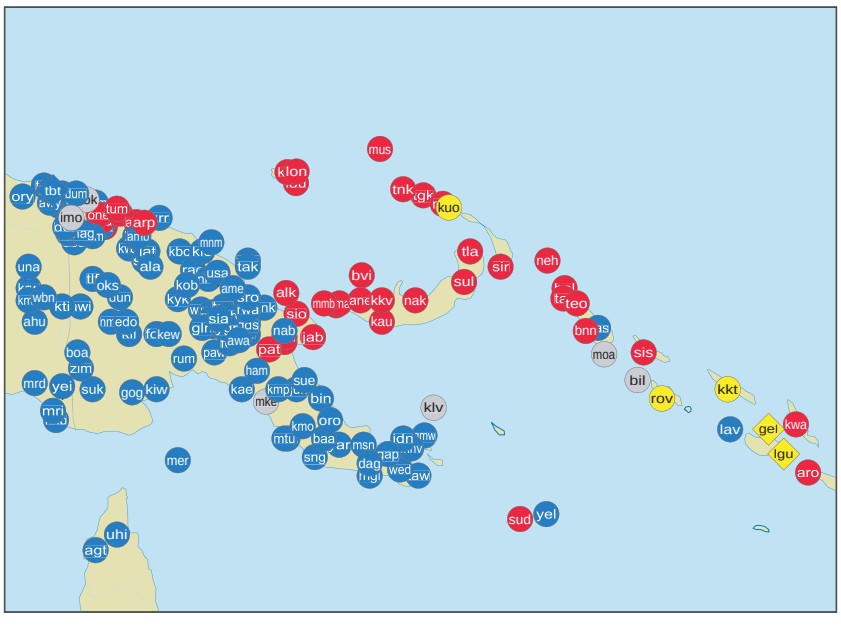
81 Order of Subject, Object, and Verb

MATTHEW S. DRYER

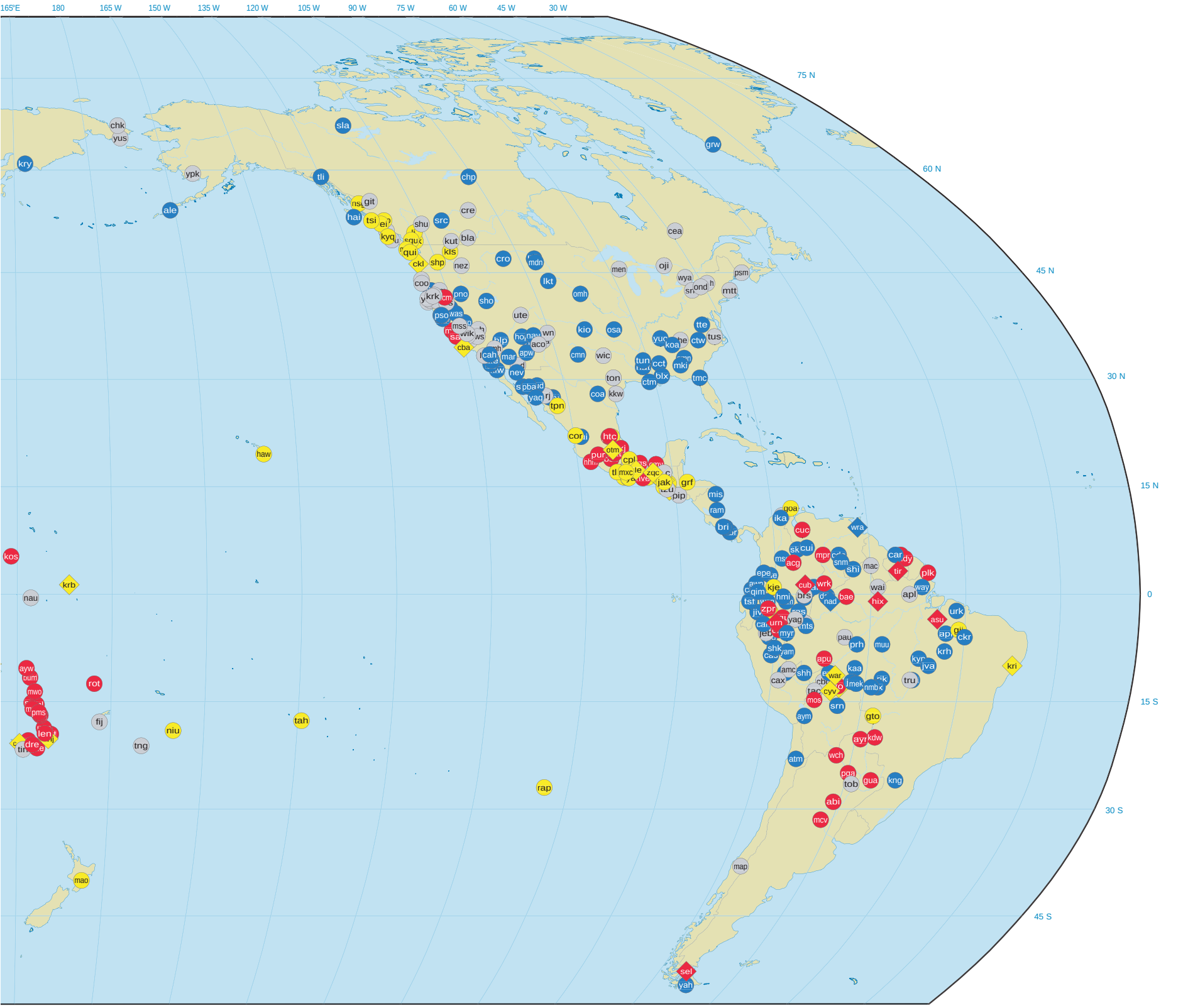


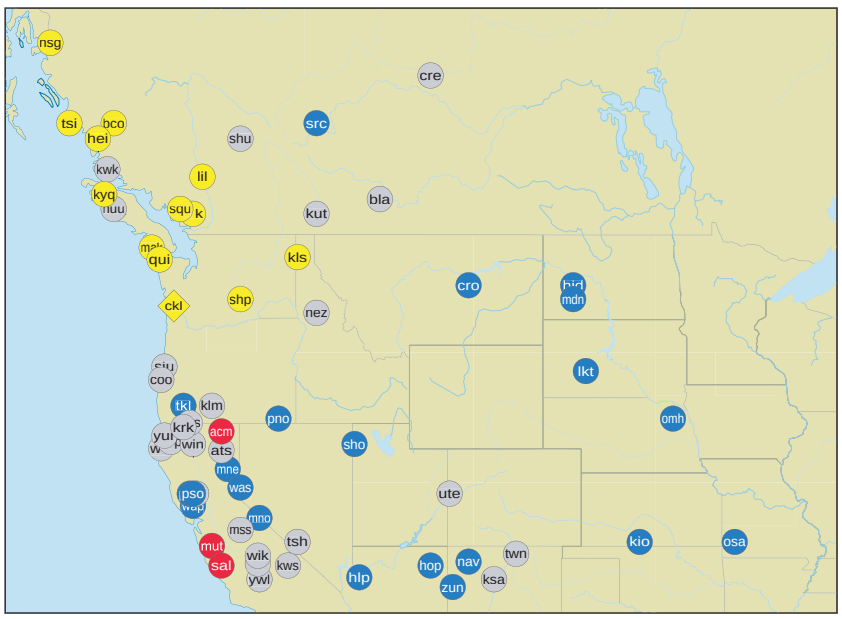


eastern Africa

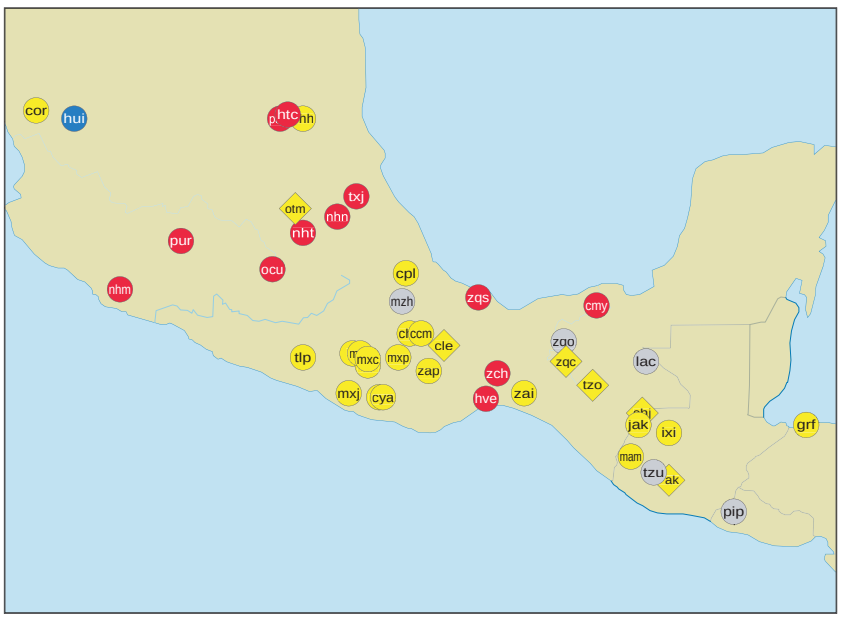


eastern New Guinea and Solomon Islands





western North America



Mesoamerica

82 Order of Subject and Verb

MATTHEW S. DRYER

1 Defining the values

This map shows the dominant order of lexical (or nonpronominal) subject and verb. The primary types are languages which are **SV (in which the subject precedes the verb)**, a type represented by English and by Turkish, as illustrated in (1a), and languages which are **VS (in which the subject follows the verb)**, exemplified by Welsh, as illustrated in (1b).

(1) a. Turkish (Kornfilt 1997: 90)
Su *kayna-dı.*
water boil-PST
S V
'The water boiled.'

b. Welsh (Williams 1980: 165)
Daeth *y* *dyn.*
come.PST.3SG the man
V S
'The man came.'

1.	Subject precedes verb (SV)	1060
2.	Subject follows verb (VS)	179
3.	Both orders with neither order dominant	105
	total	1344

As with the other maps showing word orders, there are two ways in which a language can count as SV. In some languages, SV is the only order permitted by the grammar. In other languages, both orders are grammatical, but SV order is dominant (see the box section "Determining Dominant Word Order" on p. 371). Languages in which **both orders occur and in which neither order is dominant** are classified as the third type on the map.

In languages in which the position of subjects in intransitive clauses differs from that in transitive clauses, the map shows the order in intransitive clauses. The question of what should be considered a subject in different languages and the question of whether all languages have subjects have generated considerable discussion in the literature over the years. However, these theoretical questions are largely irrelevant to this map: a language is considered SV if the single lexical argument in an intransitive verbal clause more commonly precedes the verb, but VS if such an argument more commonly follows the verb. Even languages for which people question the relevance of a notion of subject can be classified according to this criterion.

To a large extent, the distribution of SV on this map corresponds to three types on Map 81, namely SOV, SVO, and OSV. Conversely, VS on this map largely corresponds to the three types VSO, VOS, and OVS. There are a variety of ways in which these correspondences can fail to obtain. One is that there are some languages which are shown as lacking a dominant order on Map 81, because both orders of object and verb are common, but which are shown as SV on this map. German and Hungarian are languages of this sort. Another way in which these correspondences can fail to obtain is that there are languages in which both VSO and VOS order are common, and which are thus shown as lacking a dominant order on Map 81, but in which VS is dominant. Boumaa Fijian (Austronesian; Dixon 1988: 243) is an instance of such a language (see example in Chapter 83).

Another major source of differences is the fact that where subjects of transitive clauses and subjects of intransitive clauses differ in their position, this map shows the word order in intransitive clauses, while Map 81 shows word order in transitive clauses. For the vast majority of languages, the position of subjects is the same in intransitive clauses as in transitive clauses. However, there are a number of types of languages in which it is not.

The first type are languages in which word order follows an ergative pattern (see Chapters 98–100), and here there are two subtypes. First, there are languages which can be described as Ergative-Verb-Absolutive. An example of such a language is Muna (Austronesian; eastern Indonesia), illustrated in (2).

(2) Muna (van den Berg 1989: 150, 163)

a. *no-tende* *tora* *dahu*
3SG.REALIS-run again dog
V S
'The dog ran again.'

b. *o* *katogha* *ne-mbolaku* *kenta* *topa*
ART crow 3SG.REALIS-steal fish dry
S V O
'The crow stole dried fish.'

Languages like Muna are shown as VS on this map, but as SVO on Map 81. Languages of this type are common in Mesoamerica; examples include Huave (isolate; Mexico; Stairs and Hollenbach 1981: 335), Huastec (Mayan; Mexico; Edmonson 1988: 565), Tepehua (Totonacan; Mexico; Watters 1988: 12), and Michoacán Nahuatl (Uto-Aztecan; Mexico; Sischo 1977: 313). This type is also represented by Iaai (Austronesian; New Caledonia; Tryon 1968: 62–3) and Paumarí (Arauan; Brazil; Chapman and Derbyshire 1991: 163–6).

There are also languages which are the mirror image of Muna, in which the order can be described as Absolutive-Verb-Ergative: these languages are shown as OVS on Map 81 and as SV on this map. In fact, three of the nine OVS languages shown on Map 81 are of this type: Päri (Nilotic; Sudan; Andersen 1988), Mangarrayi (Mangarrayi; northern Australia; Merlan 1982), and Ungarinjin (Wororan; northwestern Australia; Rumsey 1982).

In some languages whose word order can be roughly described as ergatively based, the situation is often more complex than this characterization implies. For example, in Muna (illustrated above in (2)), subjects more often precede intransitive verbs if there is a prepositional phrase in the clause, as in (3).

(3) Muna (van den Berg 1989: 152)
kenta *topa* *no-ndawu-mo* *ne* *wite*
fish dry 3SG.REALIS-fall-PERF LOC earth
S V PP
'The dried fish fell to the ground.'

Hence the word order is sensitive to the presence of another nominal element (noun phrase or prepositional phrase) in the clause. Somewhat similarly, Salinan (Hokan; California) employs SVO order in transitive clauses containing a lexical subject and object and VS order in intransitive clauses, but again the pattern is not really an ergative one, in that transitive subjects also follow the verb if the object is pronominal and is realized only by the verb morphology (Dryer 1989). Hence, in Salinan, the order of subject and verb depends, not on the transitivity of the clause, but on whether there is a lexical object in the clause. Despite these complexities, both Muna and Salinan are shown on Map 81 as SVO and on this map as VS.

There are also languages in which there are positional differences between intransitive subjects and transitive subjects, but where the difference is not so categorical. These include languages in which the dominant order for transitive subjects is SV, but where neither order is dominant for intransitive subjects. This type is found among a number of languages of Europe (e.g. Spanish, Bulgarian, Latvian). There are also languages in which intransitive subjects more commonly follow the verb, but in which neither order is

dominant for transitive subjects. Examples of such languages include Domari (Macalister 1914), and Ostuacan Zoque (Mixe-Zoquean; Mexico; Engel and Longacre 1963: 335–6).

Among languages in which both orders of intransitive subject and verb are common and neither is dominant, a few exhibit what can be described as a split intransitive pattern, where more agentive intransitive subjects pattern with transitive subjects in preceding the verb while less agentive intransitive subjects pattern with objects in following the verb. This is illustrated in (4) for Arawak (Arawakan; Suriname); (4a) illustrates SVO order in a transitive clause; (4b) illustrates SV order in an intransitive clause with a more agentive subject; and (4c) illustrates VS order in an intransitive clause with a less agentive subject.

(4) Arawak (Pet 1987: 108, 109, 161)

a. *miaka* *aba* *wadili* *sika* *khali* *da-myn*
yesterday INDEF man give cassava.bread me-to
'Yesterday, a man gave cassava bread to me.'

b. *li* *wadili* *osa* *bahy-nro* *miaka*
the man go home-to yesterday
'The man went home yesterday.'

c. *alekhebe-ka* *li* *wadili*
happy-IND the man
'The man is/was happy.'

The split in Arawak is more specifically governed by the stativity of the verb, rather than by some factor like volitionality: subjects of verbs denoting events precede the verb while subjects of verbs denoting states follow the verb. Thus the postverbal subjects occur with verbs that correspond semantically to adjectives in English. In Mokilese (Oceanic; Micronesia), subjects of intransitive verbs denoting events obligatorily precede the verb, while subjects of intransitive verbs denoting states may either precede or follow the verb (Harrison and Albert 1976: 299–300).

Note that while this map is based on the order of subject and verb for intransitive clauses in the case of languages where it is known that there is a difference between transitive and intransitive clauses, some sources do not discuss differences between transitive and intransitive subjects. It is possible that the behaviour of intransitive subjects in some languages is sufficiently different that further data would show that they are not coded correctly on this map.

2 Geographical distribution

SV order is clearly much more common than VS order and is widely found all over the world. But it is worth noting two areas which are overwhelmingly SV, in which both VS languages and languages lacking a dominant order of subject and verb are infrequent. One of these is mainland Eurasia, where the exceptions are primarily in Europe and in eastern Siberia. The other is New Guinea. Because VS is less common, I will concentrate here on describing where VS order is found.

Despite the instances noted above in which transitive and intransitive subjects differ in their position, the overall geographical distribution of VS order is rather similar to that of VSO and VOS languages on Map 81. For most of the ways in which they differ, the differences involve languages which are shown as VS languages on this map, but not as VSO or VOS on Map 81, so that there are more VS languages on this map than languages which are shown as VSO or VOS on Map 81. There are two areas in Africa in which VS order is common: one is an area in eastern Africa extending up into Sudan, representing languages in various branches of Eastern Sudanic within Nilo-Saharan, including a number of Nilotic languages; the other is in North Africa, represented by a number of Berber languages. It should be noted that in earlier times, VS order was more widespread in North Africa and the Middle East, being the dominant order in many ancient Semitic languages and in Ancient Egyptian (see Map 81A). There is a small pocket of VS Celtic languages in northwestern Europe; otherwise VS order is not attested in Europe or mainland Asia. VS order is widespread among Austronesian languages, though mixing with SV in a complex pattern. The areas among Austronesian languages in which VS is common include: (1) an area in and around the Philippines, extending north to Taiwan and south to the northern part of Borneo; (2) northern Sumatra and islands off the west coast of Sumatra extending north to include the non-Austronesian Nicobarese languages (which are Mon-Khmer); and (3) most Polynesian languages across the Pacific. I am aware of only one VS language among all the non-Austronesian ("Papuan") languages of the New Guinea area: Kuot, spoken on New Ireland (Chung and Chung 1996: 1–4). There are a few VS languages in Australia, mainly in the south-east. VS order is very common in two well-defined linguistic areas in North America: the Pacific North-West, including the northwestern corner of the continental United States and the southwestern corner of Canada; and Mesoamerica, extending from southern Mexico into Guatemala, though not all languages of Mesoamerica are VS. It is scattered around South America, with the clearest concentration found in and around Bolivia.

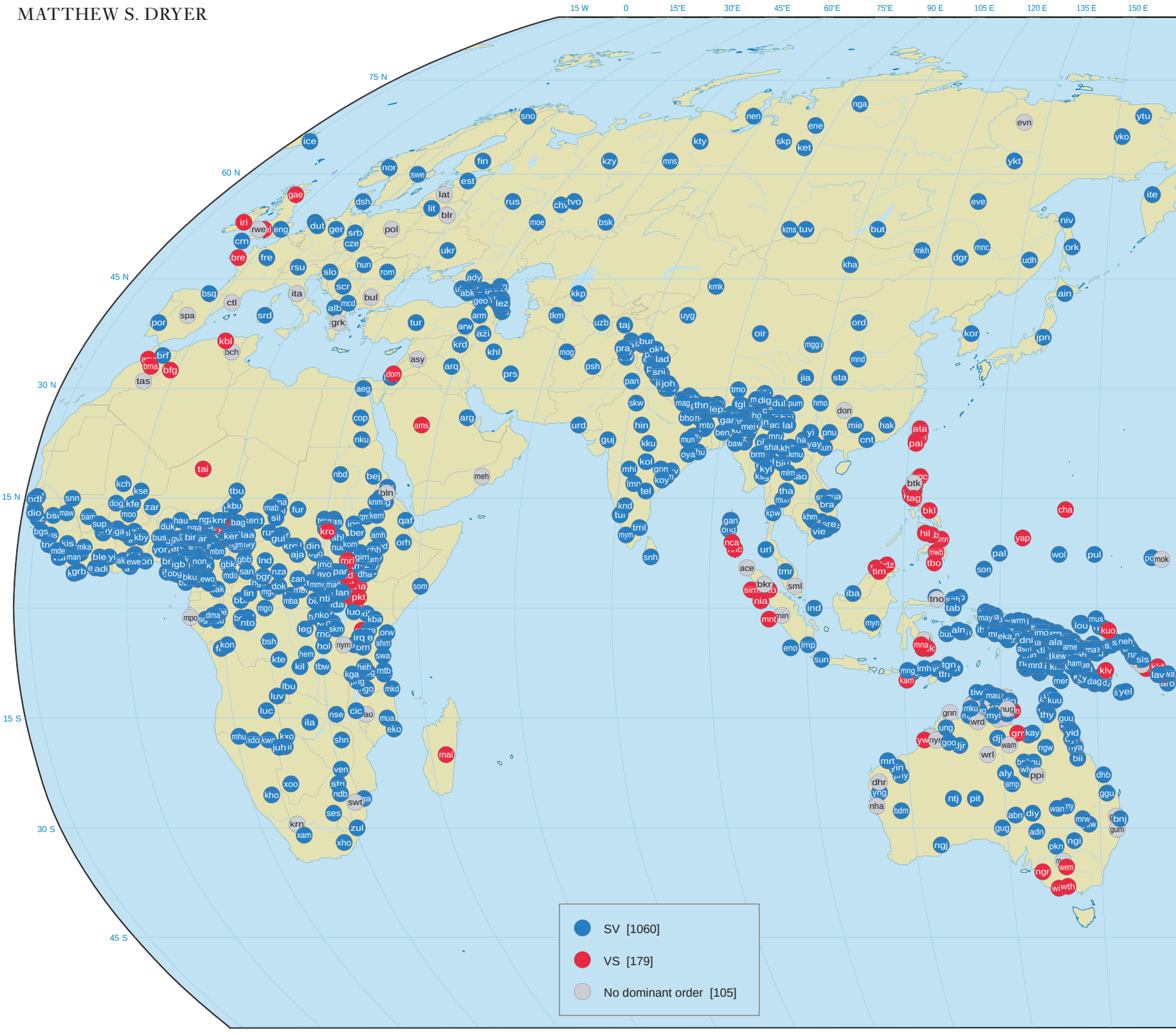
There are two areas in which languages tend to lack a dominant order on this map that are roughly coterminous with areas in which languages tend to lack a dominant order on Maps 81 and 83: these are Australia and North America. But there is one small area in which a number of languages are shown as lacking a dominant order on this map, where a number of these languages are shown as having a dominant order on Map 81: this area is around the Mediterranean, including both Indo-European languages in Europe and Berber languages in Africa. It is also worth noting that languages lacking a dominant order of subject and verb are more common around regions in which VS order is common than around regions in which SV is common.

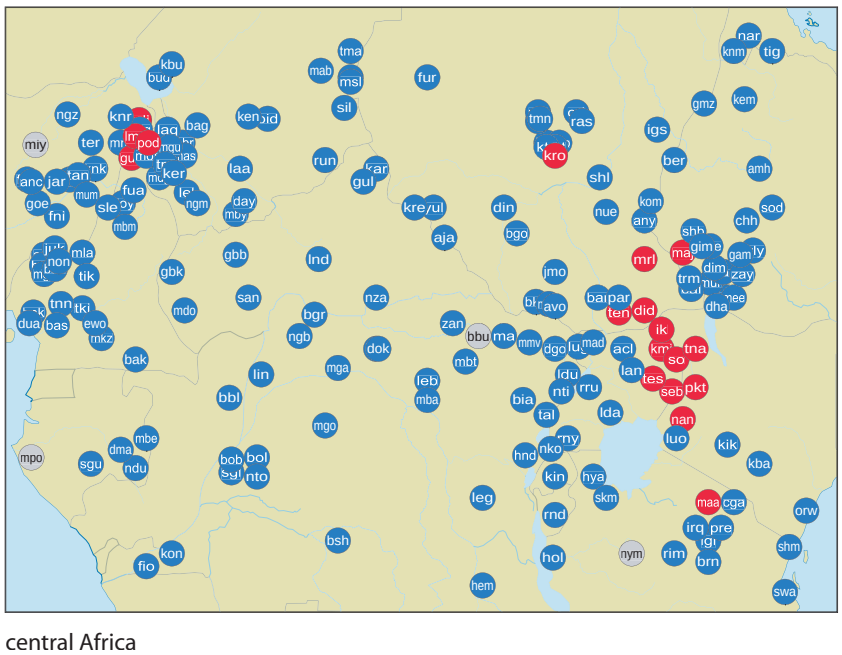
3 Theoretical issues

One theoretical question that arises is why SV order is so much more common than VS. A popular answer to this question is that it reflects a tendency for subjects to be topics in some sense (Tomlin 1986). In fact, as noted above, there are two subtypes of SV languages shown on this map: those which are fairly rigidly SV, and those with flexible order in which SV order is more common. One might surmise that in languages of the latter sort, SV order is more common because of a tendency for subjects to be topics, and that this is simply grammaticalized in languages of the former sort. However, evidence in the tradition of Givón (1983) shows that many languages seem to place what can be described as topical elements late in clauses, casting doubt on the notion that there is a universal pragmatic preference for topics to occur early in sentences. The question of why SV is preferred thus has not yet been satisfactorily answered.

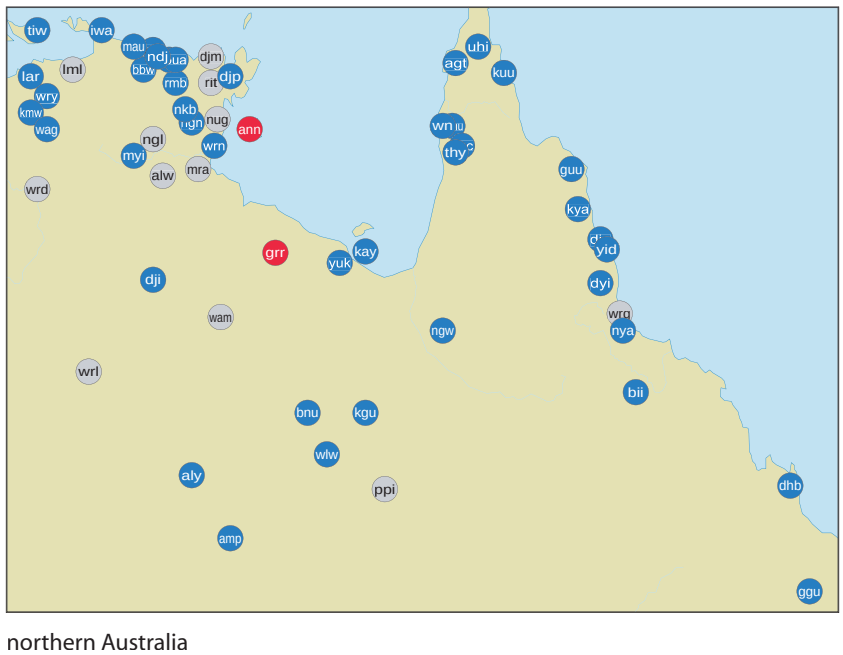
82 Order of Subject and Verb

MATTHEW S. DRYER

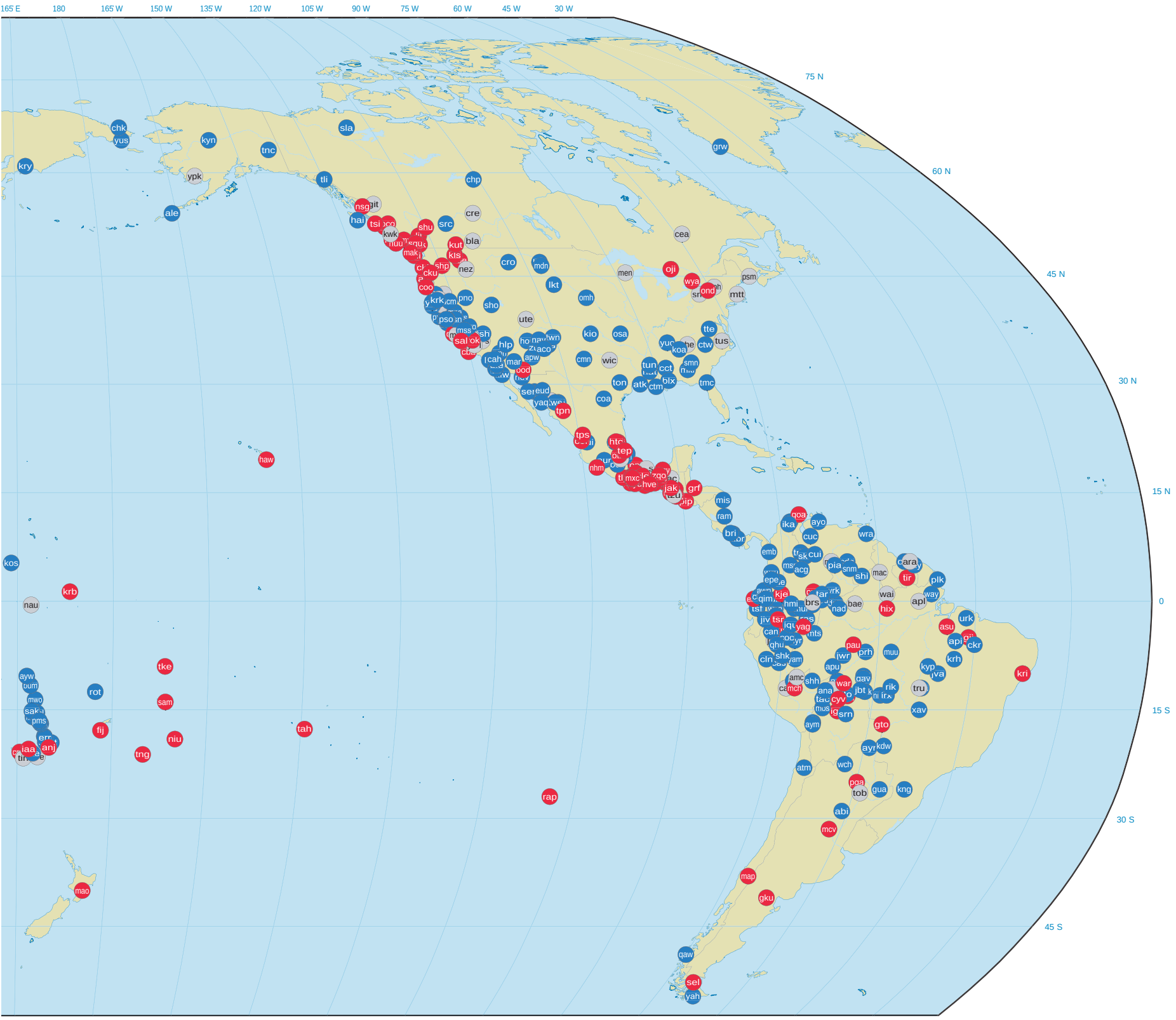


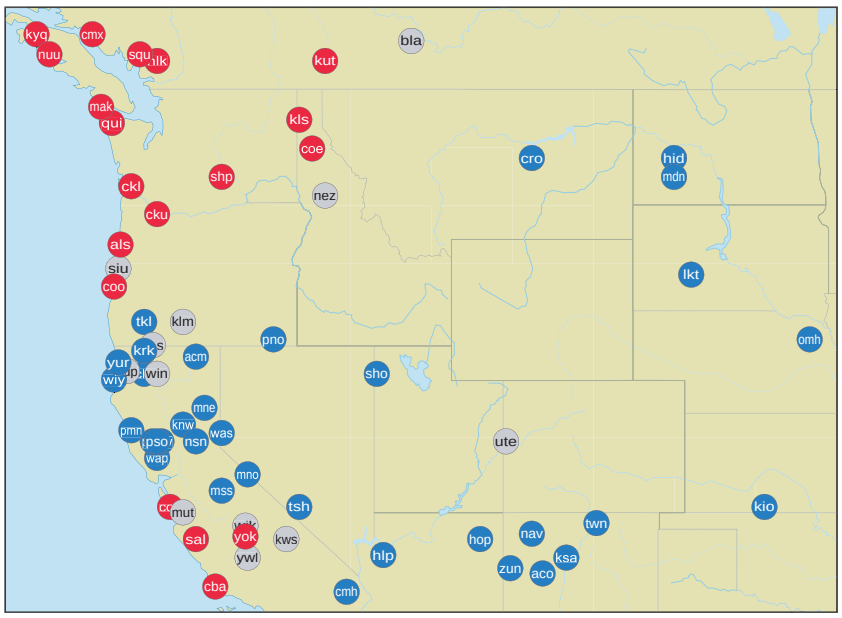


central Africa

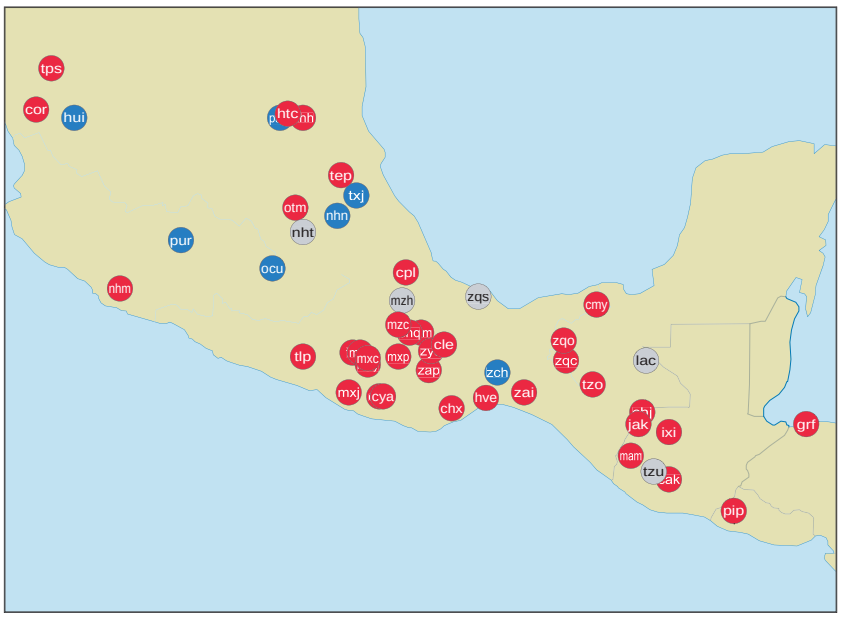


northern Australia





western North America



Mesoamerica

83 Order of Object and Verb

MATTHEW S. DRYER

1 Defining the values

This map shows the dominant order of lexical (nonpronominal) object and verb. As with Map 81, the notion of object is defined semantically, as the P or most patient-like argument (see discussion under Map 81) in a transitive clause. The primary types shown are languages which are **OV (in which the object precedes the verb)**, illustrated by Turkish in (1a), and languages which are **VO (in which the verb precedes the object)**, illustrated in (1b) by Gulf Arabic, the variety of colloquial Arabic spoken in Kuwait, Bahrain, Qatar, the United Arab Emirates, and eastern Saudi Arabia.

(1) a. Turkish (Underhill 1976: 51)
Mehmed-i *gör-dü-m.*
Mehmet-ACC see-PST-1SG
O V
'I saw Mehmet.'

b. Gulf Arabic (Holes 1990: 119)
ʔakalaw *sandwiich-aat*
eat.3PL sandwich-PL
V O
'They ate sandwiches.'

1.	Object precedes verb (OV)	640
2.	Object follows verb (VO)	639
3.	Both orders with neither order dominant	91
	total	1370

The third type is languages with **both orders with neither order dominant**; see the box section "Determining Dominant Word Order" on p. 371. A number of different subtypes of this type are discussed below. Note that the map does not distinguish languages in which only one order is possible and languages in which both orders are possible but one is dominant.

The map restricts attention to lexical noun phrases, ones consisting of a noun (plus possible modifiers), rather than objects consisting of just a pronoun. In some languages, pronominal objects occur in a different position from lexical objects. For example, in French, in which lexical objects normally follow the verb, pronominal objects normally precede the verb, as in (2).

(2) French
Je *le* *vois.*
I him see
'I see him.'

Because lexical objects normally follow the verb in French, it is shown on the map as VO.

To a large extent, the SOV type shown on Map 81 corresponds to the OV type on this map. There are two other types on Map 81 that are OV, namely OVS and OSV, but these types are quite rare. Conversely, there are three types on Map 81 that correspond roughly to VO on this map, namely SVO, VSO, and VOS. There are a number of ways, however, in which these correspondences are not exact.

First, there are a number of languages which are shown as languages lacking a dominant order on Map 81, but which are classifiable as OV or as VO on this map. Some of these are languages in which one order of object and verb is dominant but which lack a dominant order of subject and verb in transitive clauses. The most common subtype of such languages consists of languages in which SVO is a common order in transitive clauses, but where VSO or VOS (or both) is also common. Syrian Arabic is an example of a language of this sort (Cowell 1964: 407, 411).

There are also languages in which OV is the dominant order and in which both SOV and OVS order are common so that they lack a dominant order on Map 81. Macushi (Cariban; Brazil) is fairly rigidly OV, but SOV and OVS occur with about the same frequency (Abbott 1991: 25). In addition, there are languages in which the frequency of the two orders of object and verb depends on whether there is a lexical subject in the clause. For example, in Tonkawa (isolate; Texas), both SOV and SVO are common in clauses with both a lexical subject and a lexical object; but OV order is much more common in clauses lacking a lexical subject (based on my own text counts of texts in Hoijer 1972). Similarly, Yukulta (Tangkic; Queensland, Australia) is shown as SVO on Map 81, but as OV on this map, since OV is reported to be preferred if there is no lexical subject while SVO order is preferred when there is a lexical subject (Keen 1983: 229).

There are also languages for which there is a dominant order both for the order of object and verb and for the order of subject and verb, but which do not have a dominant order for subject, object, and verb. Among these are languages where both VSO and VOS order are common but neither can be considered dominant. An example of such a language is Boumaa Fijian (Austronesian), illustrated in (3); only the context will determine which noun phrase is subject in a clause of the form Verb+NP+NP.

(3) Boumaa Fijian (Dixon 1988: 243)
e *rai-ca* *a* *gone* *a* *qase*
3SG see-TR ART child ART old.person
'The old person saw the child.' or
'The child saw the old person.'

There are also many languages shown on this map that are not shown on Map 81. These are languages for which it is clear from the available materials that the language is OV or that it is VO, but where the materials do not provide enough information to determine its type for Map 81. Most languages shown on Map 81 with a specific word order (i.e. those not shown as lacking a dominant order) are shown on this map either as OV or as VO. An example of an exception is Paakantyi (Pama-Nyungan; New South Wales, Australia), which is SVO in clauses containing a lexical subject and a lexical object, but in which both OV and VO are common in clauses lacking a lexical subject (Hercus 1982: 236).

Languages in which neither OV nor VO is dominant fall into two sorts. On the one hand, there are languages with flexible word order where both orders are common and the choice is determined by extragrammatical factors. Many Australian languages, such as Ngandi (Gunwinyguan; Northern Territory, Australia; Heath 1978), are examples of this. A second class of language in which both OV and VO are common are languages in which word order is primarily determined syntactically, but in which there are competing OV and VO constructions. German is an instance of this, in that VO order is used in main clauses in which there is no auxiliary verb, as in (4a), while OV order is used in clauses with an auxiliary verb, as in (4b), and in subordinate clauses introduced by a subordinator, as in (4c).

(4) German
a. *Anna* *trink-t* *Wasser.*
Anna drink-3SG water
V O
'Anna is drinking water.'

b. *Anna* *ha-t* *Wasser* *getrunken.*
Anna have-3SG water drink.PST.PTCP
O V
'Anna has drunk water.'

c. *Hans* *sag-t,* *dass* *Anna* *Wasser* *trink-t.*
Hans say-3SG that Anna water drink-3SG
O V
'Hans says that Anna is drinking water.'

A number of languages in Africa are similar to German in employing OV order in clauses containing auxiliaries, but VO order in clauses lacking an auxiliary. The example in (5) illustrates this for Kisi (Atlantic, Niger-Congo; Guinea): (5a), without an auxiliary verb, is SVO, while in (5b), with the present progressive auxiliary *có*, the verb follows the object.

(5) Kisi (Childs 1995: 249, 250)

a. *kɛ̀ùwó* *lɔ̀wá* *sàá*
snake bite Saa
'The snake bit Saa.'

b. *Fàlà* *có* *lɛ́ɛ́ŋndó* *yìkpàá*
Fallah PRES.PROG machete sharpen
'Fallah is sharpening the machete.'

Other instances in Africa, but far to the east of Kisi, include Nuer (Western Nilotic; Sudan; Crazzolara 1933), Dinka (Western Nilotic; Sudan; Nebel 1948), and Dongo (Ubangian, Niger-Congo; Democratic Republic of Congo; Tucker and Bryan 1966: 131).

Other instances of languages with syntactically determined order of object and verb are a number of Central Sudanic languages in eastern Africa, including the Moru-Ma'di languages, in which there are two constructions which can be broadly characterized as perfective and imperfective (or past and non-past), in which the perfective construction is SVO, while the imperfective construction is SOV. The examples in (6) from Moru (Central Sudanic, Nilo-Saharan; Sudan) illustrate this.

(6) Moru (Tucker and Bryan 1966: 47)

a. *má=nya* *ŋgá*
1SG=eat something
'I ate something.'

b. *má* *ŋgá* *ɔ̀nya*
1SG something eat
'I am/was eating something.'

This contrast is not purely one of tense or aspect. For example, in Avokaya, another Moru-Ma'di language, infinitival phrases are invariably OV while imperative clauses are invariably VO (Kilpatrick 1981: 98).

There are also languages in which the order of object and verb is partly sensitive to speech-act type. For example, both Savi (Indic; Afghanistan; Buddruss 1967: 61–2) and Iraqw (Cushitic; Tanzania; Whiteley 1958: 64) are normally rigidly OV, but both allow VO in imperative clauses.

2 Geographical Distribution

The distribution of OV order is similar to that described for SOV order in Chapter 81. OV predominates over much of Asia, except in the south-east. It also predominates in New Guinea, the exceptions being either languages along the north coast or on islands offshore; many of these exceptions are Austronesian. In Australia, OV predominates over VO but competes with languages in which neither OV nor VO order is dominant; even among those classified here as OV, the order of object and verb is generally relatively flexible. In the Americas, OV is the dominant order outside two areas where VO predominates, Mesoamerica and the Pacific North-West. In Africa, it is found to the west, north, and north-east of the large area in which VO order is found, although the map is a bit misleading in that some of the areas in which OV order is found exhibit more genealogical diversity, so that in terms of genealogical groups, VO is less predominant in Africa than the map might suggest.

VO order is found in Europe and North Africa and among Semitic languages of the Middle East. It is the dominant type in Africa, though there are many OV languages around the periphery of the area in which VO is dominant. It is found in a large area stretching from China and South-East Asia through Indonesia, the Philippines, and the Pacific. Although it is the minority type in the Americas, there are two very well-defined areas that are almost exclusively VO, namely the Pacific North-West (western Canada and the northwestern part of the continental United States) and Mesoamerica. Elsewhere in the Americas, VO order is found in a number of Algonquian languages of eastern Canada, in a number of languages of California, and sprinkled throughout South America, particularly among the languages in the southern half.

Languages in which neither OV nor VO order is dominant are particularly common in Australia, and to a somewhat lesser extent, in North America. The Moru-Ma'di and Western Nilotic languages mentioned above, in which the choice between OV and VO is grammatically determined, form a clearly defined small area in eastern Africa.

3 Theoretical issues

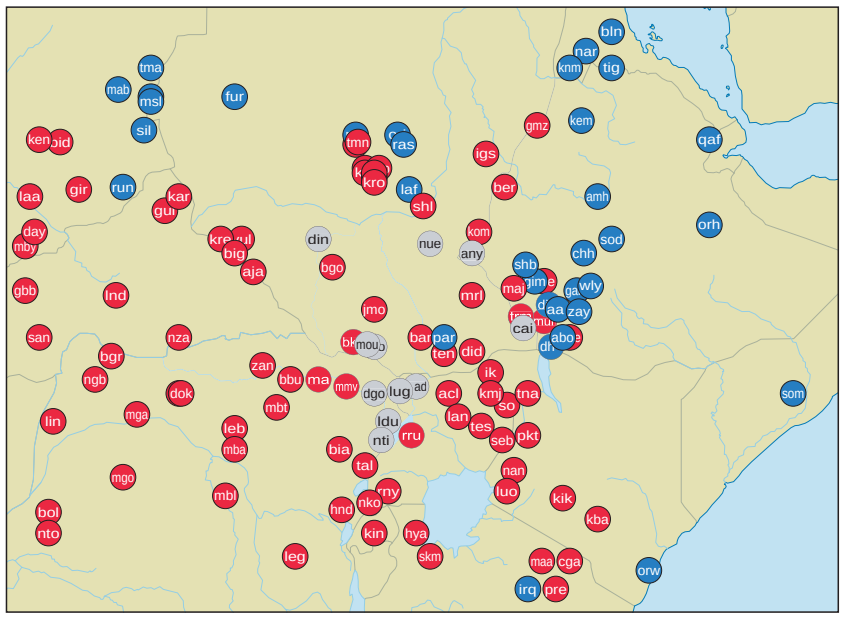
The order of object and verb has received considerable attention because of the fact that a large number of other features are predictable from it, at least in a statistical sense (Greenberg 1963, Hawkins 1983, Dryer 1992). See Chapters 95, 96, and 97 for discussion. For example, OV languages tend to be postpositional (see Chapters 85 and 95), genitive before noun (see Chapter 86), adverb before verb, complementizer at end of clause, and standard-marker–adjective order in comparative clauses, while VO languages tend to exhibit the opposite orders. The patterns are sometimes more complex than this. For example, while VO languages almost exclusively place relative clauses after nouns, both orders of relative clause and noun are common among OV languages (see Chapter 96). In addition, there are some word order features which do not correlate with the order of object and verb. For example, contrary to some claims, the order of adjective and noun does not correlate with the order of object and verb (Dryer 1988a, 1992; see Chapter 97).

While it is often assumed in the literature that the order of object and verb has some privileged status among the various pairs of elements which correlate in order with each other, this assumption has not been supported. There is really no other good candidate among the various pairs of elements for such a privileged status. Perhaps the best alternative candidate would be adposition type (prepositions versus postpositions); but many languages lack adpositions, yet still exhibit correlations among other pairs of elements. An alternative view is that no pair of elements has a privileged status; rather, there are just many pairs that correlate with each other, and the order of object and verb is just one of those pairs of elements.

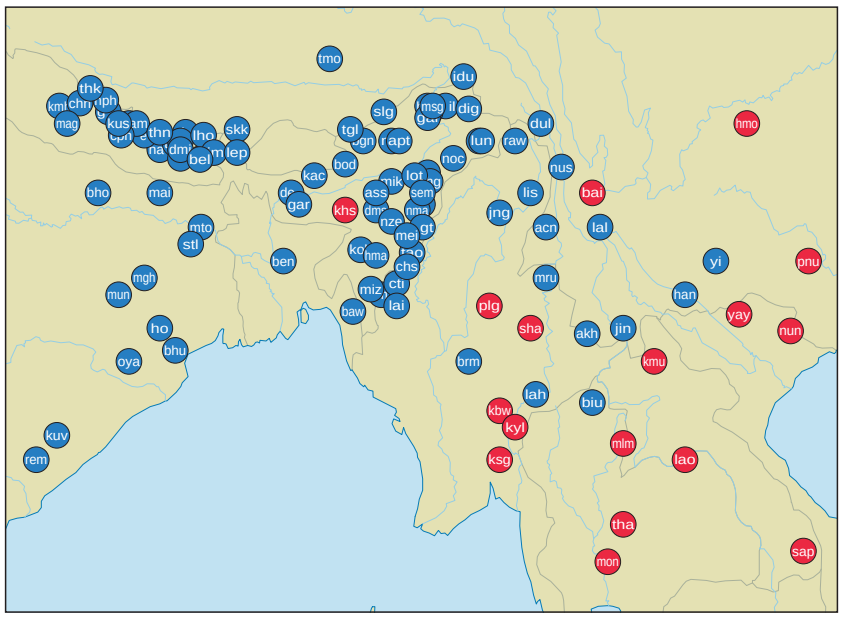
83 Order of Object and Verb

MATTHEW S. DRYER

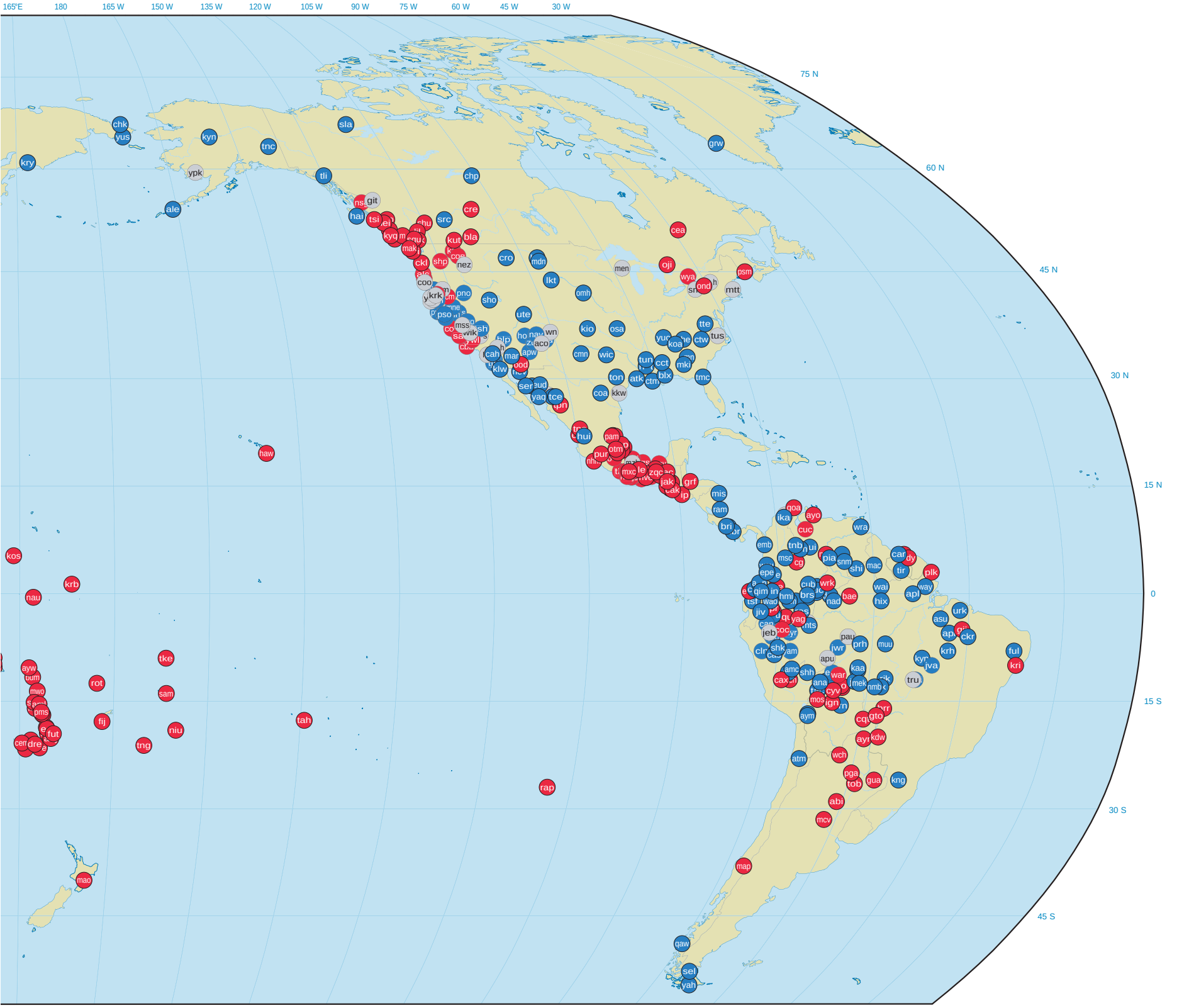
OV [640]
VO [639]
No dominant order [91]

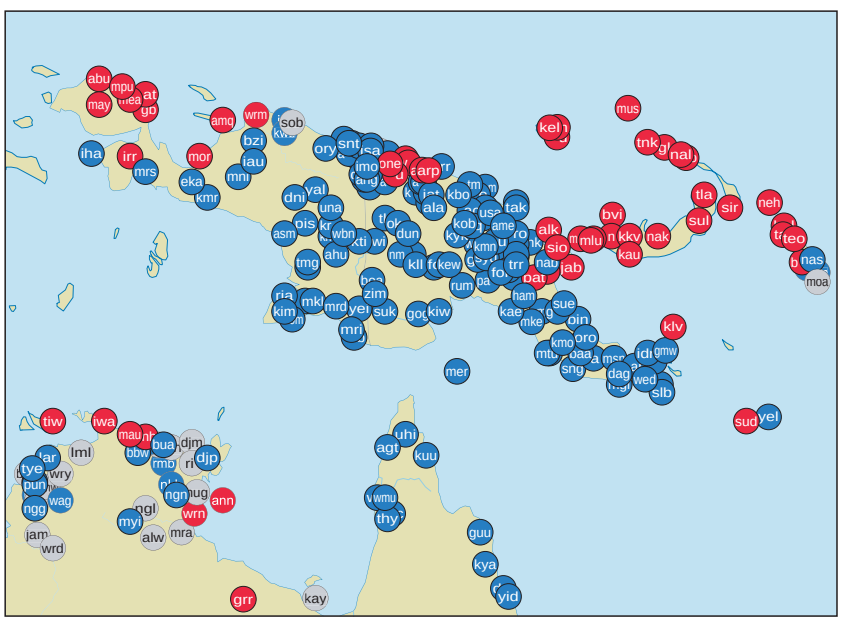


eastern Africa

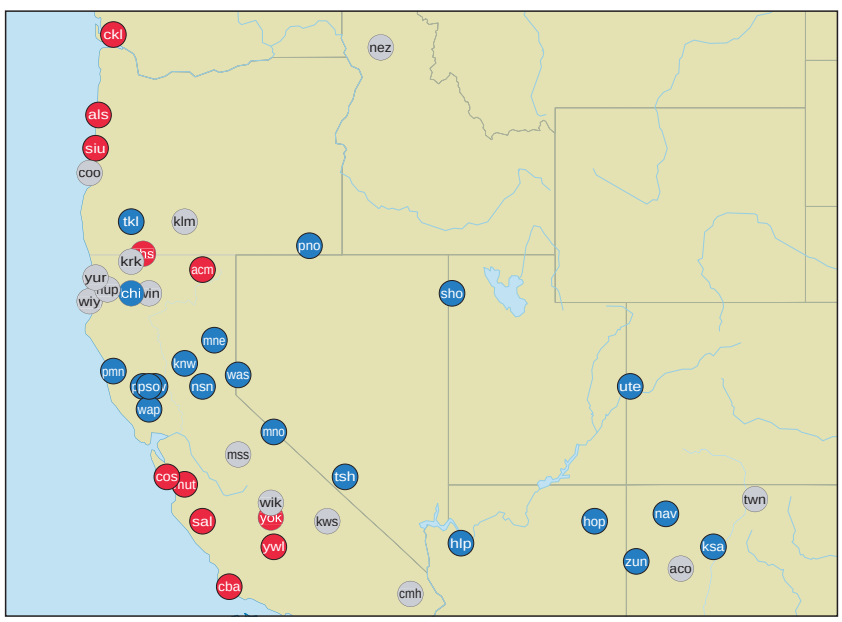


from Nepal to Vietnam





New Guinea and northern Australia



western United States

84 Order of Object, Oblique, and Verb

MATTHEW S. DRYER (with Orin D. Gensler)

1 Defining the values

This map shows the dominant order of lexical object, oblique phrase, and verb. An oblique phrase is a noun phrase or adpositional phrase (prepositional or postpositional) that functions as an adverbial modifier (or "adjunct") of the verb. Of the six logically possible orders, five are attested, as illustrated in (1), where *X* stands for an oblique phrase.

(1) a. English
Mary opened [the door] [with a key].
V O X

b. Mandarin (Li and Thompson 1981: 390)
tāmen [zài fángzi-hòumian] xiūli diànshìjī
they at house-behind repair television
X V O
'They repair televisions behind their house.'

c. Slave (Athapaskan; northern Canada; Rice 1989: 997)
t'eere [denǫ gha] ʔerákeeʔee wihsį
girl REFL.mother for parka 3.made
X O V
'The girl made a parka for her mother.'

d. Nagatman (isolate; Papua New Guinea; Campbell and Campbell 1987: 8)
na [mo me] [ke na] hohui-në-taya
and fish OBJ these with look.for-1.SUBJ-3PL.OBJ
O X V
'And we look for fish with these.'

e. Kairiru (Oceanic; Papua New Guinea; Wivell 1981: 151)
ei porri tamiok a-pik [gege-i nat nai]
3SG axe 3SG-take [from-3SG child that]
O V X
'He/she took the axe from that child.'

The sixth logical possibility, VXO, is not attested. There are many languages in which this order is possible, but there is no language in the sample in which it occurs as the dominant order (although it is standard in Irish with pronominal objects; Ó Siadhail 1989: 207). The map also shows languages whose word order is sufficiently flexible that there is no dominant order for these three elements (see the box section "Determining Dominant Word Order" on p. 371). Some of these are languages with highly flexible order that are also shown on Map 81 as ones lacking a dominant order for the order of subject, object, and verb. Others are ones that have a dominant order on Map 81, but not on this map. There are a number of languages which are shown as SOV on Map 81, but which are shown as lacking a dominant order on this map because the object and oblique are apparently freely ordered with respect to each other. An example of such a language is Lower Grand Valley Dani (Trans-New Guinea; Papua, Indonesia; Bromley 1981).

1.	Verb-object-oblique order (VOX)	189
2.	Oblique-verb-object order (XVO)	3
3.	Oblique-object-verb order (XOV)	45
4.	Object-oblique-verb order (OXV)	23
5.	Object-verb-oblique order (OVX)	37
6.	More than one order with none dominant	152
	total	449

The XVO type illustrated in (1b) from Mandarin is quite rare; the only known instances are varieties of Chinese. The rarity of this type and the nonattested status of VXO mean that VO languages are overwhelmingly VOX. But all three types of OV languages are widely attested. Among them, XOV is the most frequent on the map (45 languages), OVX next most frequent (37 languages), and OXV least frequent (23 languages). But this seriously overrepresents the true frequency of OVX, which is relatively rare in the world. The frequency on the map is a methodological artifact, and emerges from the fact that many grammars state the ordering of O and V and of X and V but not of O and X. In such cases, if the language is of the type OV&XV, nothing can be concluded about the ordering of O and X, and so such languages had to be omitted from the map. By contrast, if the language is of the type OV&VX, then the ordering OX follows automatically, and so the language is always included on the map. Hence the overrepresentation. If we include languages that are OV and XV but for which we were not able to determine a dominant order of X and O or for which both orders of X and O are common, we find that out of 318 OV languages, 281 are XV and only 37 are VX.

Oblique phrases include phrases expressing location (such as source and goal), instruments (as in (1a)), benefactives (as in (1c)), and comitatives (as in English *with Pat* in *I played tennis with Pat*). Indirect objects (recipients with ditransitive verbs) are not included; in some languages they pattern with the direct object, in others with obliques. Temporal expressions are not included, since in many languages they exhibit positional properties different from other obliques, often occurring at the beginning of the clause.

In English, oblique phrases are normally expressed by prepositional phrases. In many other languages, however, they do not involve adpositions, but rather case affixes on the head noun (and sometimes on modifiers as well), as illustrated by the instrumental suffix in (2) from Yidiny (Pama-Nyungan; Queensland, Australia).

(2) Yidiny (Dixon 1977: 263)
ŋaɲḏi duguɽ balga:l ḏirga:-da
1PL.SUBJ house.ABS make.PST blady.grass-INSTR
'We made a house with blady grass.'

In some languages, oblique phrases occur or can occur without an adposition or case affix. In the example from Blackfoot (Algonquian; Alberta and Montana) in (3), the only morphemes corresponding to the prepositions *from* and *with* in the English translations are the morphemes in the verb glossed 'SOURCE' and 'ACCOMP'; the verb itself means 'to arrive from somewhere with someone'.

(3) Blackfoot (Frantz 1991: 106)
nít-ohpok-oht-o'too-maw amá
1-ACCOMP-SOURCE-arrive-DIRECT.3SG this.3SG
nit-ohkíímaan-a Omahk-oyis-i
1POSS-wife-ANIM.SG big-lodge-OBV
'I arrived from Edmonton with my wife.'

For the purposes of this map, phrases like *on the table* in English *I put the book on the table*, which can be viewed as arguments of the verb, are treated as obliques, primarily because in most languages their order with respect to the verb and object is the same as clear instances of obliques. An exception is Awa Pit (Barbacoan; Colombia), in which purely optional obliques precede the object, as in (4a), while oblique arguments occur between the object and verb, as in (4b).

(4) Awa Pit (Curnow 1997: 61, 102)
a. *awa=na tɨ=kasa piya waa-mtu*
person=TOP stick=with corn sow-IMPF.PTCP
X O V
'The Awa sow corn with a stick.'

b. *na=na pala tɨm=ki win-ta-w*
1SG NOM=TOP plantain basket=at put-PST-LOC.SUBJ
O X V
'I put the plantains in the basket.'

Because purely optional obliques are presumably more frequent than oblique arguments, Awa Pit is shown on the map as SXOV. There are other languages in which different types of obliques exhibit different positional preferences relative to the object. For example, in Suena (Trans-New Guinea; East Papua New Guinea), which is SOV, temporal and locative expressions tend to occur between the subject and the object, while instrumental expressions tend to occur between the object and the verb (Wilson 1974: 98); Suena is shown on the map as a language lacking a dominant order for these elements.

2 Geographical distribution

Since most VO languages are VOX, the distribution of VOX languages is similar to that described in Chapter 83 for VO (though there are fewer languages shown on this map, since data on this characteristic is available for fewer languages). As noted above, XVO order is restricted to varieties of Chinese. All three types of OV languages (XOV, OXV, and OVX) are fairly widely distributed. The distribution of OXV versus XOV order does not exhibit any particularly clear geographical pattern. OVX order, however, exhibits stronger areality. It is most common in West Africa, where it is the universal or near-universal pattern in several language groups, notably Mande and Senufo Gur, and in individual languages such as Koyraboro Senni, illustrated below. In several other West African groups, notably Kru, it occurs in alternation with VOX depending on whether an auxiliary is present (see below). It is also found in a number of languages in Australia and in Central America and adjacent areas in Colombia.

3 Theoretical issues

The typology that forms the basis of this map has rarely been discussed. The two most frequent types, VOX and its mirror image XOV, share two properties: (1) the oblique phrase occurs on the same side of the verb as the object; and (2) the object occurs adjacent to the verb. The fact that oblique phrases occur on the same side of the verb as the object is often taken to be something so obvious as to be in no need of explanation. It is true that both are dependents of the verb, but so are subjects, and, except for Chinese, oblique phrases occur on the same side of the verb as the subject only in languages in which the object does so as well. That the object most often occurs adjacent to the verb was observed by Tomlin (1986) and conforms to general principles whereby grammatical complements should occur closer to their heads than adjuncts. But the fact that there are many OV languages which are OXV or in which objects and obliques are freely ordered with respect to each other shows that such principles governing complements and adjuncts are only tendencies.

One question is why we find the asymmetry we find between VO and OV languages: while VO languages are overwhelmingly VOX, all three types of OV languages (XOV, OXV, OVX) are widely attested. There are two questions here. The first is why OXV languages are common while their mirror image, VXO, is unattested. The second is why it is that most languages in which the oblique phrase occurs on the opposite side of the verb from the object are OVX, rather than XVO. Neither question has an obvious answer in the literature.

A curious property of OVX languages is that many of them are SAuxOVX, placing auxiliary elements, whether verbal or nonverbal, between the subject and the object (Gensler 1994, 1997). While this phenomenon is apparently restricted to Africa, it is found in widely separated areas. Language groups in West Africa with SAuxOVX as their standard word order include Mande, Senufo (Gur), as illustrated by Supyire (Senufo; Mali) in (5a), and Koyraboro Senni (Songhay; Mali), as in (5b).

(5) a. Supyire (Carlson 1994: 274)

u	*sí*	*sìɲciiyí*	*cya*	*mìì*	*á*
3SG	FUT	firewood.DEF	seek	1SG	for
	Aux	O	V	X	

'She will fetch firewood for me.'

b. Koyraboro Senni (Heath 1999a: 139)

ay	*ga*	*nooru*	*wiri*	*[ay*	*baaba*	*ga]*
1SG	IMPF	money	seek	1SG	father	on

'I will seek money from my father.'

This pattern is also found in Tunen, a Bantu language which is atypically SOV and which is spoken in Cameroon, quite distant from the OVX languages in West Africa, as illustrated in (6).

(6) Tunen (Dugast 1971: 309)

mɛ̀	*ná*	*mìɔkɔ*	*bokòn*	*[ò*	*nuìy]*
1SG	PST	stones	raised	LOC	river

'I raised the stones from the river.'

The presence of SAuxOVX in distant branches of Niger-Congo suggests that this order may be reconstructible for Niger-Congo. Plausibly, the S-TAM-O-V morphological order (where *TAM* stands for 'tense-aspect-mood') which is characteristic of the verb in most Bantu languages represents a grammaticalization of the syntactic SAuxOV pattern, further supporting the reconstruction.

This special status of SAuxOVX also extends to a number of other languages in Africa in which OVX alternates with some other word order. In various language groups, notably Niger-Congo languages of West Africa, SAuxOVX alternates with auxiliary-less SVO. This is found in Ewe and in Kru, as illustrated in (7) for Grebo (Kru; Liberia).

(7) Grebo (Innes 1966: 102, 103)

a.

ne	*du-da*	*bla*	*ne*
I	pound-DIST.PST	rice	PTCL

'I pounded rice.'

b.

ne	*yi*	*bla*	*du-i*
I	IMPF	rice	pound-IMPF

'I am pounding rice.'

A more complex case is found in Moru-Ma'di languages (Central Sudanic, Nilo-Saharan; Uganda and Democratic Republic of the Congo). Here, we find a two-way opposition conditioned by syntactic factors and tense-aspect: auxiliary-less SVO versus S(Aux)OVX. This is illustrated in (8) for Ma'di (Uganda), (8a) illustrating the AuxOV order, (8b), the OVX order.

(8) Ma'di (Blackings and Fabb 2003: 441, 364)

a.

má	*rɨ̀*	*dʒɔ́*	*`sī-rē*
1SG	PROG	house	NONPST.build-SUBORD

'I was/have been building a house.'

b.

ɔ́pɨ́	*ká*	*bɔ̀ŋgú*	*dʒè*	*sàbúnì*	*sɨ̀*
Opi	3SUBJ	clothes	wash	soap	SOURCE

'Opi is washing clothes with soap.'

Similar facts obtain in Lendu and Ngiti, languages belonging to a different branch of Central Sudanic from the Moru-Ma'di languages, but spoken close by. Similarly, in Dinka and Nuer (Nilotic; Sudan) the word order varies with a number of factors, but in main clauses containing an auxiliary, the dominant order is SAuxOVX.

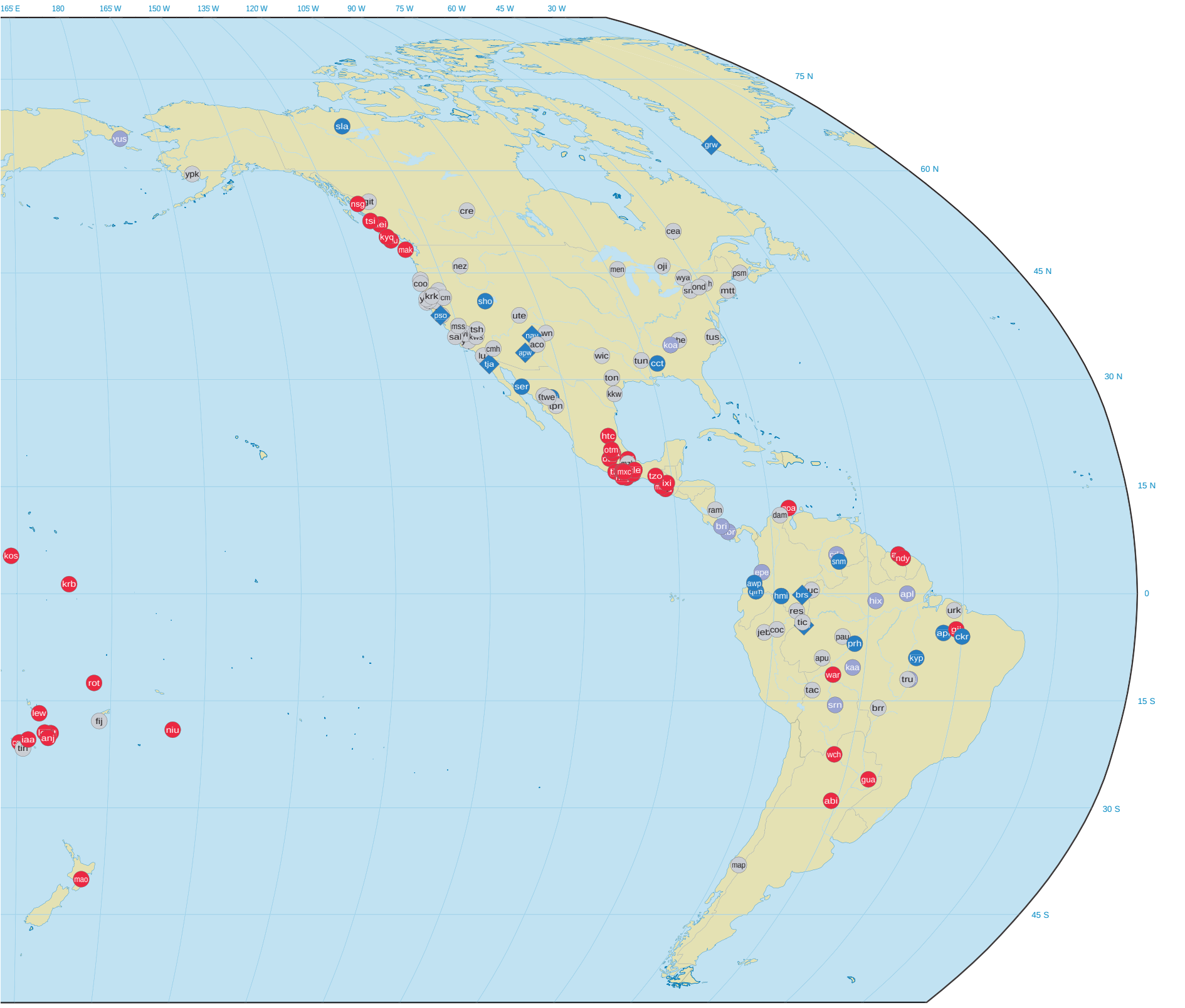
84 Order of Object, Oblique, and Verb

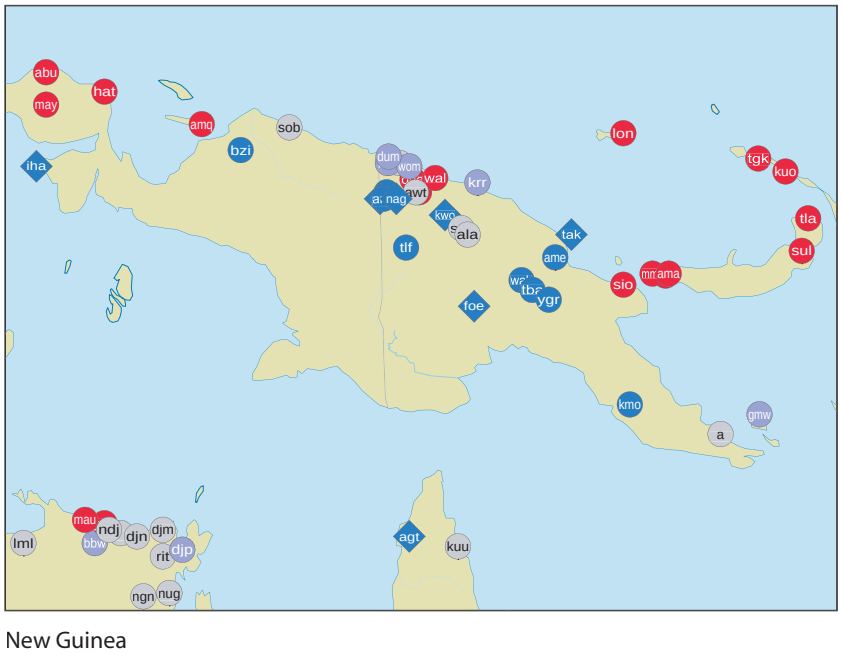
MATTHEW S. DRYER (with Orin D. Gensler)

VOX [189]
XVO [3]
XOV [45]
OXV [23]
OVX [37]
No dominant order [152]

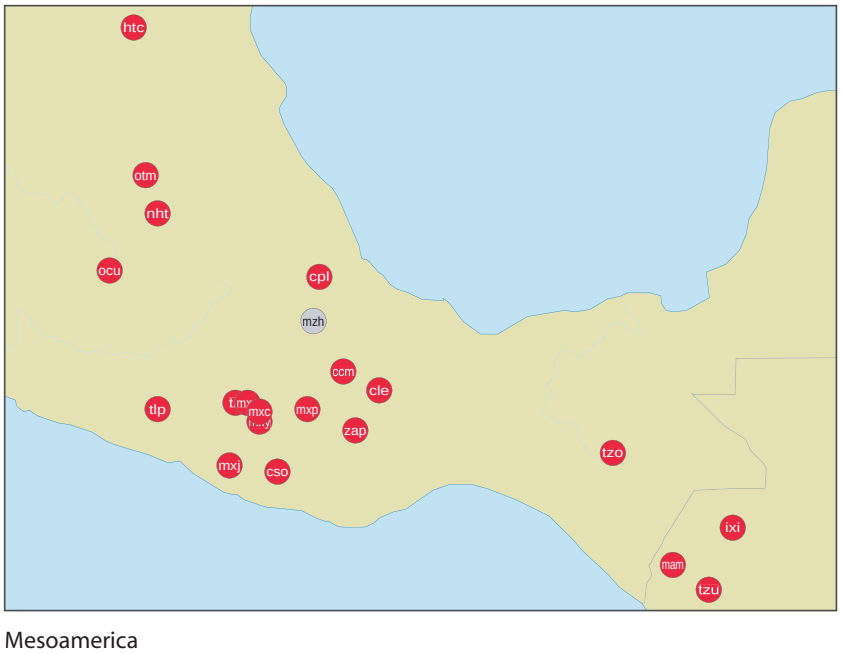
Nigeria and surrounding areas

eastern Africa





New Guinea



Mesoamerica

85 Order of Adposition and Noun Phrase

MATTHEW S. DRYER

1 Defining the values

This map shows the order of adposition and noun phrase. The two primary types of adpositions are **prepositions** and **postpositions**: prepositions precede the noun phrase they occur with, as in English and in the Boumaa Fijian (Austronesian) example in (1a), while postpositions follow the noun phrase they occur with, as in the Lezgian (Nakh-Daghestanian; Russia) example in (1b).

(1) a. Boumaa Fijian (Dixon 1988: 216)

au	*na*	*talai*	*Elia*	***i***	*'Orovou*
1SG	FUT	send	Elia	**to**	'Orovou
				Prep	NP

'I'll send Elia to 'Orovou.'

b. Lezgian (Haspelmath 1993: 218)

duxtur-r-in	***patariw***	*fe-na*
doctor-PL-GEN	**to**	go-AOR
NP	Postp	

'She went to doctors.'

1.	Postpositions	520
2.	Prepositions	467
3.	Inpositions	7
4.	More than one adposition type with none dominant	52
5.	No adpositions	28
	total	1074

A word is treated here as an adposition (preposition or postposition) if it combines with a noun phrase and indicates the grammatical or semantic relationship of that noun phrase to the verb in the clause. Some languages also employ adpositions to indicate a relationship of a noun phrase to a noun (especially in a genitive/possessive relationship); however, if the only candidates in a language for adpositions are in the genitive construction, they are not treated as adpositions here.

In some languages, some or all of the functions of adpositions are carried by **case affixes** on nouns, as in the example in (2) from Ngalakan (Gunwinyguan; Northern Territory, Australia).

(2) Ngalakan (Merlan 1983: 46)

ŋañjuḷa-ŋini-ʔwala	*ŋu-yerk-gaŋiñ*
eye-1SG.POSS-**from**	1SG.3SG-come.out-CAUS.PST.PUNCT

'I removed it from my eye.'

Case affixes and adpositions can be referred to together as **case markers**. While some linguists occasionally apply the terms *preposition* or *postposition* to case affixes with meanings corresponding to prepositions in European languages, case affixes are not treated as adpositions on this map (see Map 51 on case affixes). On the other hand, many languages have case markers which are not separate words phonologically but whose position is still determined syntactically. The most common instances of this are case markers that are clitics which attach phonologically to the first or last word in the noun phrase, as illustrated by the postpositional clitic in (3) from Kunuz Nubian (Nilo-Saharan; Egypt)

(3) Kunuz Nubian (Abdel-Hafiz 1988: 283)

[esey	*kursel]=lo*	*uski-takki-s-i*
[village	old]=LOC	born-PASS-PST-1SG

'I was born in an old village.'

Such clitic case markers, which attach to modifiers of the noun if they are at the beginning or end of the noun phrase, are treated here as instances of adpositions since they combine syntactically with noun phrases, even though they are not separate phonological words. A number of languages in which modifiers always precede the noun, and in which the case marker always occurs at the end of the noun phrase (and hence on the noun), are in principle ambiguous as to whether the case marker should be treated as a case suffix or a postpositional clitic; for the purposes of this map, I treat such case markers as case suffixes and not as postpositional clitics.

The map also shows a rare third type of adposition, what I will call **inpositions**, adpositions which occur or can occur inside the noun phrase they accompany. In Anindilyakwa (isolate; Northern Territory, Australia) the inpositions are second-position clitics within the noun phrase, attaching phonologically to the end of the first word in the noun phrase, as in (4), in which the inposition attaches to the word for 'small' in the noun phrase meaning 'small stick'.

(4) Anindilyakwa (Groote Eylandt Linguistics-langwa 1993: 202)

. . . narri-ng-akbilyang-uma	*[eyukwujiya=**manja***	*eka]*
. . . NC_1.PL-NC_2-stick.to.end-TA	[small=LOC	stick]

'. . . they stuck them (the feathers) to a little stick.'

In Tümpisa Shoshone (Uto-Aztecan; California), the inpositions appear immediately after the noun and before any postnominal modifiers (if there are any), as in (5).

(5) Tümpisa Shoshone (Dayley 1989b: 257)

[ohipim	***ma***	*natii'iwantü-nna]*	*tiyaitaiha*	*satü*
[cold.OBJ	**from**	mean-OBJ]	died	that

'He died from a mean cold.'

In (5), the inposition *ma* 'from' appears between the head noun *ohipim* 'cold' and its postnominal modifier *natti'iwantünna* 'mean'. Note that the inpositions in Tümpisa Shoshone govern the objective case on pronouns, nouns, and their modifiers (though the case is often null, as in the noun *ohipim* 'cold'), as shown on the postnominal modifier in (5). This shows that despite appearing inside the noun phrase, the inposition still determines the case inflection of words in that noun phrase. Only seven languages are shown on the map as having inpositions as the dominant adposition type, six of them in Australia.

Some languages have both prepositions and postpositions. While there are some languages in which specific adpositions can be used either as prepositions or as postpositions, in most languages of mixed adposition type, some of the adpositions are always prepositions while others are always postpositions. In some languages with both prepositions and postpositions, one type is considered dominant if there are considerably more adpositions of one type than the other or if there is reason to believe that one type is considerably more common in usage (see the box section "Determining Dominant Word Order" on p. 371). In Koyra Chiini (Songhay; Mali), for example, there are more than twice as many postpositions as prepositions and the prepositions have more specialized meanings (like 'without'), while some of the postpositions have fairly basic meanings, suggesting that postpositions are probably much more common in usage (Heath 1999b: 103, 108). If neither type can be considered dominant, then the language is shown on the map as **more than one adposition type with none dominant**, though this also includes rare instances of languages with both postpositions and inpositions, such as Hanis Coos (Oregon Coast family; Frachtenberg 1922b). For example, Koromfe (Gur, Niger-Congo; Burkina Faso and Mali) has only two prepositions and an unclear number of postpositions (but greater than two); however, one of the prepositions has very broad meaning and appears to occur with high frequency, so Koromfe is treated as a language lacking a dominant adposition type (Rennison 1997: 73, 77, 294).

The final type consists of languages which **do not have adpositions**, or at least appear not to. Some such languages only employ case affixes as case markers, as in Yidiny (Pama-Nyungan; Queensland, Australia; Dixon 1977); others lack case markers altogether, as in Kutenai (isolate; western North America). This type is underrepresented on the map because grammars do not generally say if a language lacks adpositions and one can only infer the absence of adpositions from a thorough grammar. Some languages only have one minor adposition; such languages are coded here according to the type of that one adposition. For example, Wardaman (Yangmanic; Northern Territory, Australia) has only one postposition, meaning 'like', as in (6).

(6) Wardaman (Merlan 1994: 99)
mernden *marrajbi* *ya-wurr-yanggan*
white.ABS like 3.SUBJ-3NSG.OBJ-go.POTENTIAL
'They have to be like white people.'

The words analysed here as prepositions or postpositions are often referred to by authors of grammars by some other label. For example, what are treated here as clitic postpositions are often referred to as *case suffixes* in descriptions of languages, and many grammars do not mention that the so-called case suffixes attach to modifiers of the noun rather than to the noun if the modifier is the last element in the noun phrase. Even among adpositions which are not clitics, the words that count here as adpositions are often referred to by some other term. In some grammars, for example, they are called *relators* (e.g. Derbyshire 1985 in reference to postpositions in Hixkaryana). In many languages, the words treated here as adpositions share grammatical properties with nouns or verbs and are often for that reason referred to in grammars as nouns or verbs. These shared properties generally reflect the fact that it is common for nouns and verbs to grammaticalize as adpositions, while often still retaining grammatical properties reflecting their grammaticalization source. For example, in Jakaltek (Mayan; Guatemala), prepositions inflect for their object with the same set of pronominal prefixes that indicate possessors on nouns, and the structure of prepositional phrases is the same as that of noun phrases with possessors. Thus the construction in (7a) mirrors that in (7b).

(7) Jakaltek (Craig 1977: 110, 106)
a. *y-ul* *te'* *ñah*
3-in the.CLF house
'in the house'
b. *y-ixal* *naj* *pel*
3-wife CLF Peter
'Peter's wife'

Such situations can either be described by saying that prepositions share certain properties with nouns or by saying that prepositions are a subclass of nouns. It is assumed here that the difference between these two ways of describing the situation is terminological. In fact, while Craig (1977) refers to words like *ul* 'in' in (7a) as prepositions, Day (1973: 82), in a different description of the same language, characterizes them as nouns. Thus, the fact that a set of words with adpositional meaning arguably constitute a subclass of some other class, such as nouns or verbs, is not considered here as a reason not to treat them as adpositions.

On the other hand, the fact that certain nouns (or verbs) in a language sometimes translate into prepositions in English is not sufficient grounds for them to be treated here as adpositions. There must be some reason to believe that they have grammaticalized to some extent, that they are to some extent grammatically distinct from other nouns (or verbs). For example, in languages with serial verb constructions, the equivalent of an instrumental adposition is often expressed by a verb meaning 'use', as in the example in (8) from Mandarin.

(8) Mandarin (Li and Thompson 1981: 597)
tāmen *yòng* *shǒu* *chī-fan*
3PL use hand eat-food
'They eat with their hands.'

But in the absence of evidence of grammaticalization, an example like that in (8) is not sufficient to conclude that the word *yòng* 'use' functions as anything but a normal verb. Conversely, in Maybrat (West Papuan; Papua, Indonesia), there is a word *ae* 'at' which is morphologically and syntactically a verb. The example in (9a) illustrates it functioning as a main verb, while the example in (9b) illustrates it functioning prepositionally.

(9) Maybrat (Dol 1999: 87, 88)
a. *y-ae* *Sorong*
3SG M-at Sorong
'He is in Sorong.'
b. *ait* *y-amo* *m-ae* *amah*
3SG M 3SG.M-go 3SG.F-at house
'He goes home.'

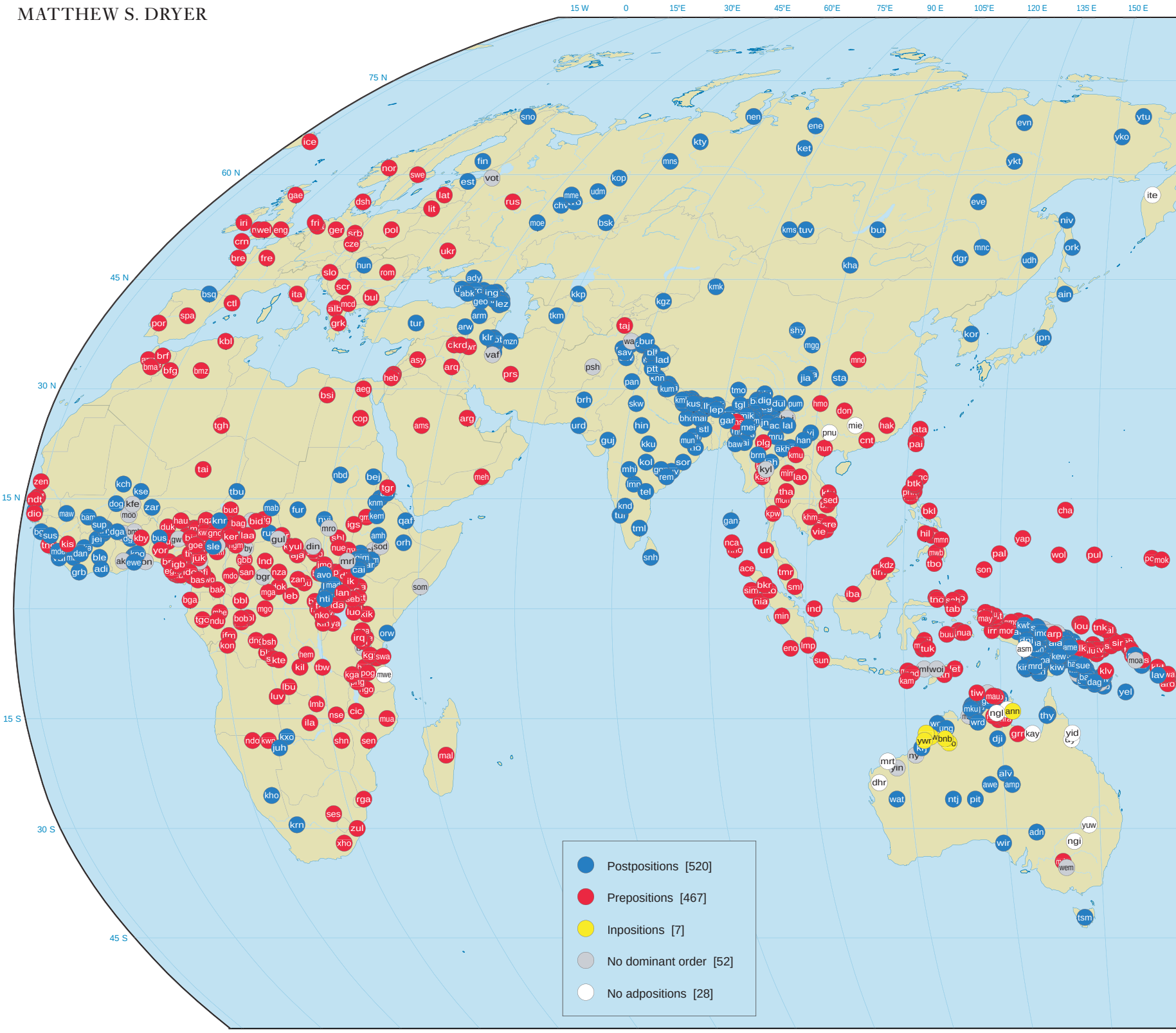
However, when *ae* is used prepositionally, as in (9b), it always occurs with a third-person singular feminine subject prefix regardless of the person, number, and gender of the subject, indicating that it is grammatically distinct from normal verbs. It thus counts as a preposition for the purposes of this map.

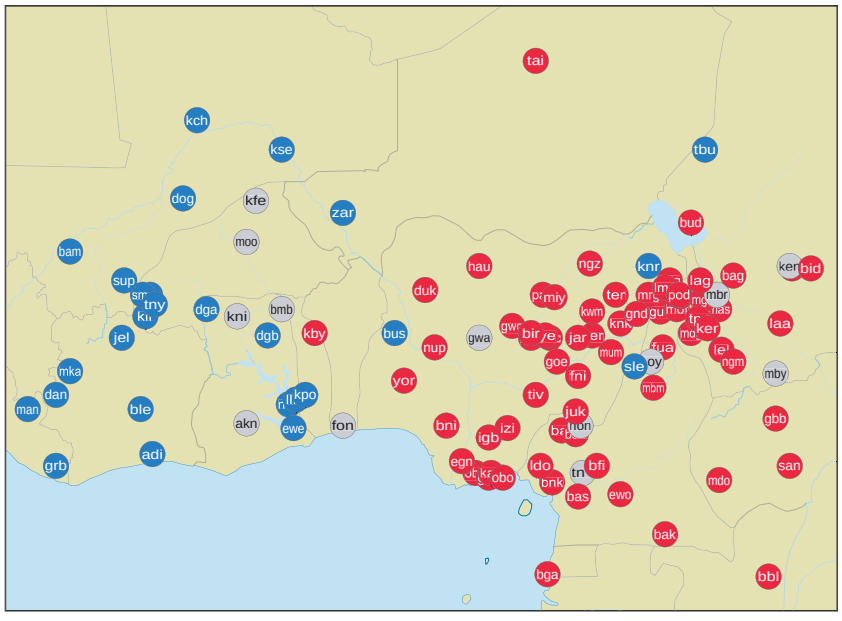
2 Geographical distribution

Because adposition type correlates strongly with the order of object and verb (see Chapter 95), the distribution of prepositions and postpositions on this map resembles the distribution of object and verb on Map 83. Prepositions predominate in the following areas: (i) Europe, North Africa, and the Middle East; (ii) central and southern Africa; (iii) a large area extending from South-East Asia through Indonesia, the Philippines, and the Pacific; (iv) the Pacific North-West in Canada and the United States; and (v) Mesoamerica. Postpositions predominate (i) in most of Asia, except in South-East Asia; (ii) in New Guinea, except in the north-west; (iii) in North America, except in the two areas noted above; and (iv) in most of South America. Postpositions are more common than prepositions in much of Australia, especially among Pama-Nyungan languages, but in the northern part of Northern Territory both types occur with comparable frequency. In fact, for many Australian languages, especially Pama-Nyungan, there is no evidence of adpositions of any sort. While prepositions predominate in Africa as a whole, there are still many languages with postpositions, including an area in West Africa and one to the north-east. There is one area in Africa stretching from Sudan and Ethiopia south-west into the northeastern corner of the Democratic Republic of the Congo where the map is quite complex. Languages with no adpositions are most common in Australia and North America, though as noted above, this type is more common than the map suggests.

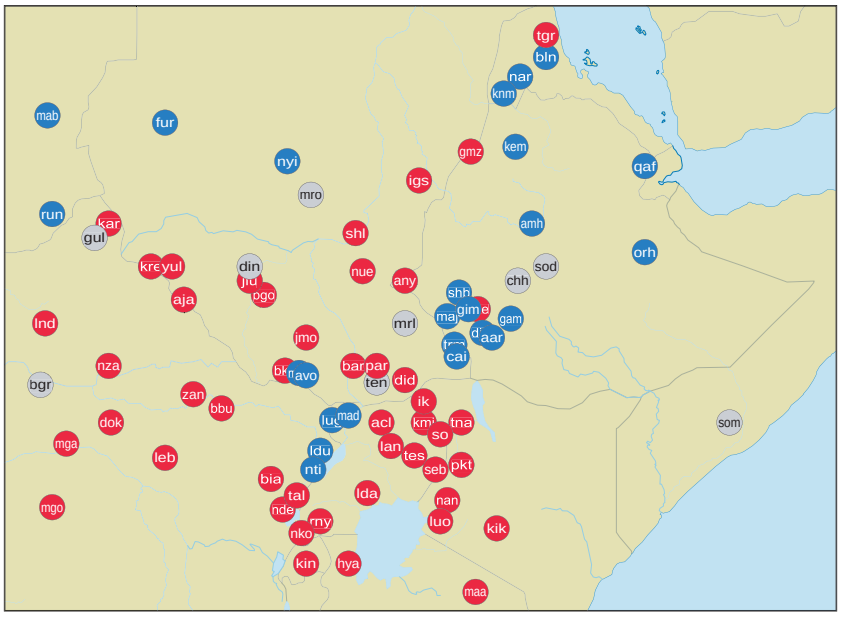
85 Order of Adposition and Noun Phrase

MATTHEW S. DRYER

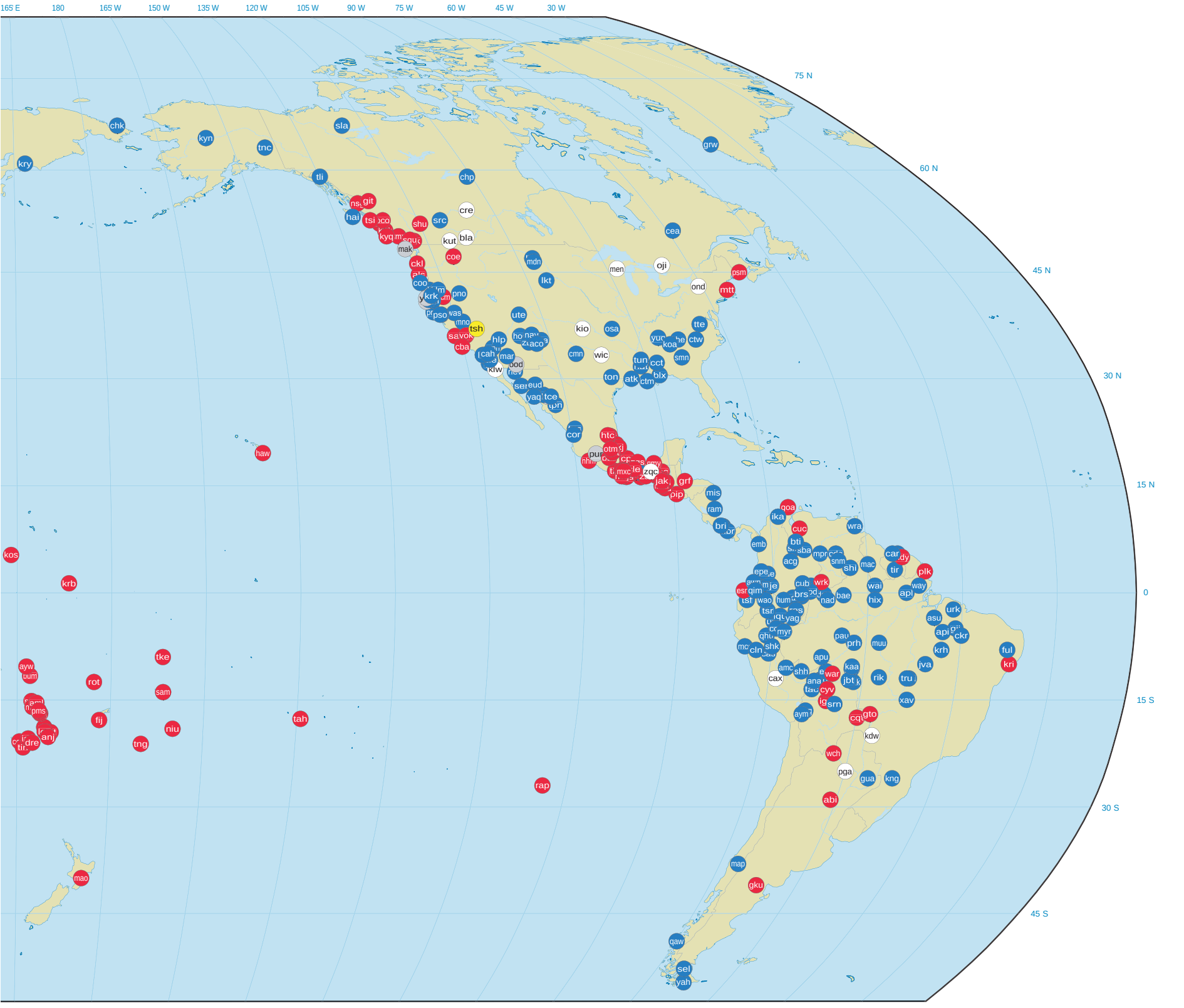


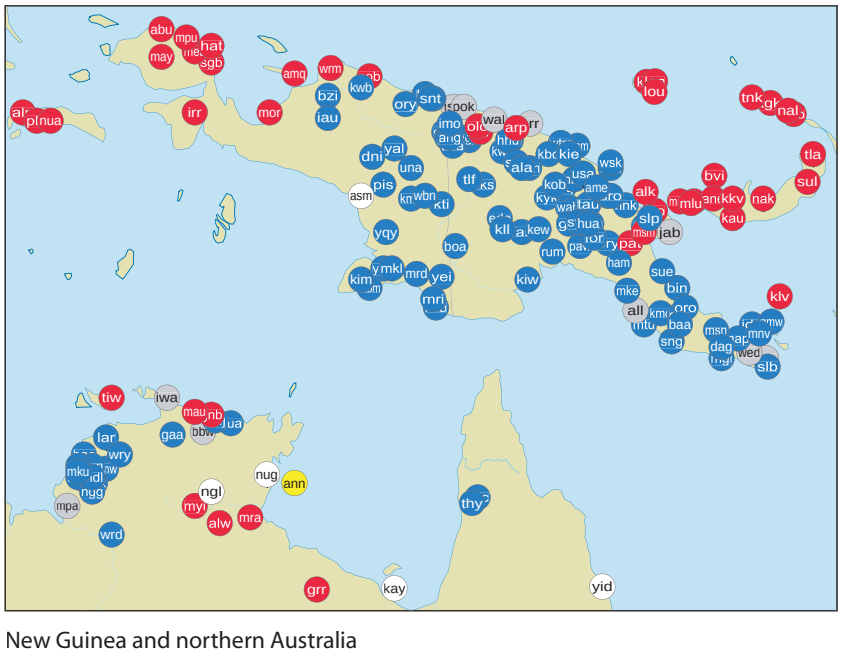


Nigeria and surrounding areas

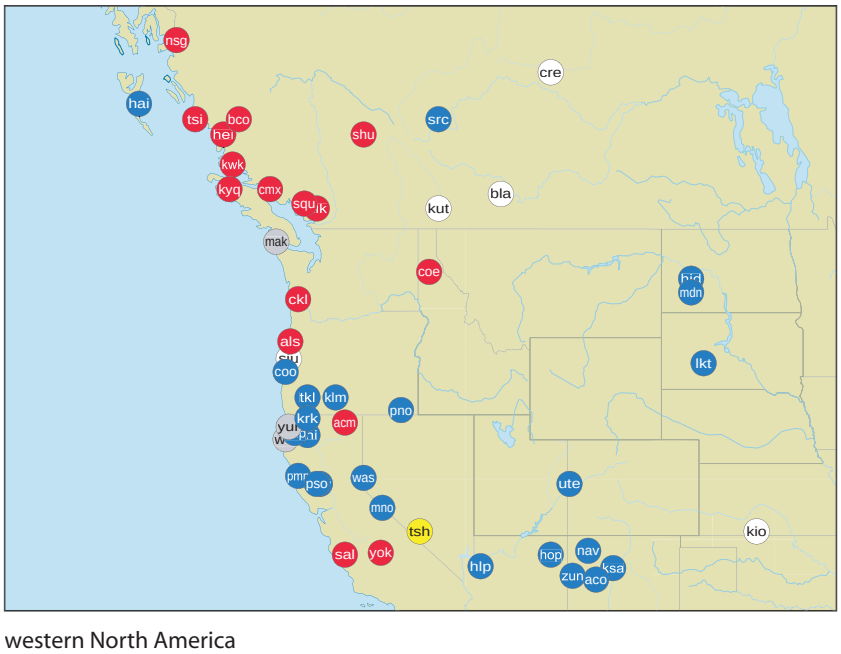


eastern Africa





New Guinea and northern Australia



western North America

86 Order of Genitive and Noun

MATTHEW S. DRYER

1 Defining the values

This map shows the order of a genitive or possessor noun phrase with respect to the head noun. In the example in (1) from Finnish, *tytön* 'the girl's' is the genitive noun phrase, while *kissa* 'cat' is the head noun, so the order in Finnish is **GenN** (genitive preceding head noun).

(1) Finnish (Sulkala and Karjalainen 1992: 227)
tytö-n *kissa*
girl-GEN cat
Gen N
'the girl's cat'

The opposite order, **NGen**, is illustrated by the example from Krongo (Kadugli; Sudan) in (2).

(2) Krongo (Reh 1985: 152)
níimò *má-Kùkkú*
mother GEN-Kukku
N Gen
'Kukku's mother'

1.	Genitive–noun (GenN)	608
2.	Noun–genitive (NGen)	415
3.	Both orders occur with neither order dominant	82
	total	1105

The genitive noun phrase is often called the **possessor** (phrase) and the head noun is sometimes called the **possessee** (noun), and the construction itself is known either as a *genitive construction* or as a *possessive construction*. It should be stressed that the term *possession* is used in this context in a broader sense than the term is used in everyday English. Two basic semantic types of genitive relations involve body parts and kinship relations, as illustrated by the French examples in (3).

(3) French
a. *le* *bras* *de* *Jean*
the arm of Jean
'Jean's arm'
b. *le* *père* *de* *Jean*
the father of Jean
'Jean's father'

But the genitive relation can involve various other semantic relations as well. In *le crayon de Jean* 'Jean's pencil', the relationship is one of possession or ownership. In *la mort de Jean* 'Jean's death', the relationship is similar to that of a subject and a verb, while in *la construction de la maison* 'the building of the house' it is more akin to the relationship of an object to a verb. Other semantic relationships are involved in *le maire de Paris* 'the mayor of Paris' and *la population de Paris* 'the population of Paris'. In fact, the genitive relation can be described as simply involving *some* semantic relation between a noun phrase and a noun, the particular relation being determined by inference from the semantics of the words involved and from the context. The order of genitive and noun in all of the above French examples is NGen.

While French uses a single construction in the examples cited above, English employs two different constructions. One, used in *John's arm* or *the governor's house*, involves the order GenN and marks the genitive noun phrase by the clitic morpheme spelled *'s*. The other, used in *the mayor of Paris* and *the population of Paris*, involves the order NGen, with the preposition *of* preceding the genitive noun phrase. The factors governing the choice between these two genitive constructions in English are complex, and involve a combination of syntactic and semantic factors (see Rosenbach 2002). Since both of these constructions are quite common in English, English is coded as a language having both orders with neither order dominant (see the box section "Determining Dominant Word Order" on p. 371). Languages in which neither order is dominant fall into two subtypes: those languages like English where there are two distinct constructions, one which is GenN and another which is NGen, and languages in which there is essentially one construction, but both orders are allowed.

Many languages distinguish two sorts of genitive constructions, one for **alienable possession** and one for **inalienable possession** (see Chapter 58). In many such languages, the order of genitive and noun is the same in the two constructions. In some languages, however, the order of genitive and noun is different for alienable and inalienable possession. For example, in Maybrat (West Papuan), inalienable possession employs the order GenN, with the head noun marked by a prefix agreeing with the genitive noun phrase in person, number, and gender, while alienable possession employs the order NGen, with the genitive noun marked with a genitive case affix. The example in (4a) illustrates GenN order with inalienable possession, while (4b) illustrates alienable possession.

(4) Maybrat (Dol 1999: 93, 97)
a. *Sely* *m-me*
Sely 3SG.F.POSS-mother
'Sely's mother'
b. *amah* *ro-Petrus*
house GEN-Petrus
'Petrus's house'

Pohnpeian (Oceanic; Ponape Island, Micronesia) exhibits the opposite pattern: NGen order is used for inalienable possession, while GenN is used for alienable possession (Rehg 1981: 165).

The map restricts attention to the position of genitive noun phrases containing nouns, rather than ones involving only a pronominal word. In many languages, the same construction is used regardless of whether the genitive is headed by a noun or is just pronominal, as in the Mandarin examples in (5): in both examples, the genitive noun or pronoun occurs first, followed by the linker particle *de*, followed by the possessed noun.

(5) Mandarin (Li and Thompson 1981: 113)
a. *rùzi* *de* *ěrduō*
rabbit LINK ear
'the rabbit's ear'
b. *wǒ* *de* *chènshān*
1SG LINK shirt
'my shirt'

In some languages, however, pronominal possession involves a distinct construction and sometimes involves a different word order. An example of such a language is French, which is NGen with genitives involving nouns, as illustrated in (3) above, but which places pronominal possessive words before the head noun, as in (6).

(6) French
son *livre*
3.POSS.M book
'his/her book'

Conversely, Tauya (Madang, Trans-New Guinea; Papua New Guinea) is GenN for nominal possessors, as in (7a), but places pronominal possessors after the head noun, as in (7b).

(7) Tauya (MacDonald 1990: 133, 131)

a. *ʔe* *fanu-na* *wate*
that man-GEN house
'that man's house'

b. *wate* *ne-pi*
house 3SG-GEN
'his/her house'

Many languages do not employ separate pronominal words, but use affixes to express such meanings, as in (8) from Kutenai (isolate; western North America).

(8) Kutenai (own data)
xaɬi-ʔis
son-3.POSS
'his/her/their son'

See Map 57 showing the distribution of such possessive affixes.

The map also excludes the position of nonreferential genitives, if a language employs a construction for these that is distinct from that used for referential genitives. Many languages are like English in employing a type of compound construction for nonreferential genitives (*deer antlers, truck tires*).

There are a number of different structures used for genitive constructions cross-linguistically, the differences between which are ignored for this map, although in some cases different orders are more commonly associated with different constructions. A crucial difference is whether either the genitive or the head noun bears affixes that are part of the genitive construction. In some languages, the genitive noun occurs with a genitive case affix, as in the example in (9) from Chukotko (Chukotko-Kamchatkan; eastern Siberia, Russia).

(9) Chukchi (Dunn 1999: 362)
ənpənacɣ-əqaj-ərɣ-en *ŋeekak*
old.person-DIM-PL-GEN daughter
'the older people's daughter'

In other languages, it is the head noun that bears an affix, typically a pronominal affix coding features of the possessor, like person, number, or gender. This is illustrated in (10) for Kutenai, where the same third-person possessor suffix *-ʔis* is used as in (8) above.

(10) Kutenai
xaɬi-ʔis *qu* *paɬkiy*
son-3.POSS that woman
'that woman's son'

In some languages, a word marking the construction as a genitive construction occurs between the genitive and the head noun. This word may be an adposition, as in the French examples in (3) above, a general linker particle used in other modifier-head constructions, as in (5) above from Mandarin, or a pronominal form, agreeing in number with the genitive noun phrase, as in (11) from Loniu (Oceanic; Manus Island, Papua New Guinea).

(11) Loniu (Hamel 1994: 94)
ɲatama *iy* *pihin*
father 3SG woman
'the woman's father'

In some languages, the genitive construction does not employ any affixes or separate words, the genitive and head noun being simply juxtaposed, as illustrated in (12) from Chalcatongo Mixtec (Oto-Manguean; Mexico).

(12) Chalcatongo Mixtec (Macaulay 1996: 108)
kačíní *peðrú*
hat Pedro
'Pedro's hat'

It is not uncommon for a language to employ two of the construction types described above, as illustrated by the examples in (4) from Maybrat.

2 Geographical distribution

Because the order of genitive and noun correlates with the order of object and verb, the general pattern on this map is similar to that of Map 83. NGen order predominates (i) in Europe; (ii) in much of Africa; (iii) in South-East Asia and among the Austronesian languages of Indonesia, the Philippines, and the Pacific; (iv) in the Pacific North-West in North America; and (v) in Mesoamerica. GenN order predominates (i) in an area of West Africa to the west of Nigeria, as well as two less well-defined areas around Ethiopia and south-west Africa; (ii) in much of Asia other than the south-east; (iii) in an area around New Guinea and extending westward into eastern Indonesia; (iv) in the Americas, except for the two areas noted above in which NGen order predominates. Both orders, as well as languages lacking a dominant order, are common in Australia, though GenN is more common.

While the map bears an overall resemblance to Map 83 showing the distribution of the order of object and verb, it differs in ways that are worth mentioning and that can be summarized by saying that GenN order often has a broader distribution than OV order, a situation that arises because many SVO languages are GenN. But because these SVO languages with GenN order tend to be adjacent to OV languages with GenN order, the specific areas in which GenN order is common are often somewhat larger than the corresponding areas where OV order is common. An example of this is the area around New Guinea: while there is a small minority of languages on the mainland of New Guinea that are VO, there are very few NGen languages (and most of these are just offshore on the north coast). Furthermore, the GenN area around New Guinea extends further west than the corresponding OV area, including the Austronesian languages of eastern Indonesia that are closest to New Guinea. Similarly, the GenN area in West Africa is much broader and more clearly defined than the corresponding OV area. Conversely, the area in South-East Asia in which VO order is found is broader than the corresponding NGen area, extending northward into China. And while VO order predominates over much of Europe, this is much less true for NGen order: Scandinavia and the Baltic are GenN and there are many languages in Europe in which neither order of genitive and noun predominates.

3 Theoretical issues

While the order of genitive and noun correlates with the order of object and verb (OV languages tending to be GenN, and VO languages NGen), it differs from other pairs of elements whose order so correlates in that SVO languages are intermediate between OV and verb-initial languages: SVO&GenN languages are as common as SVO&NGen languages.

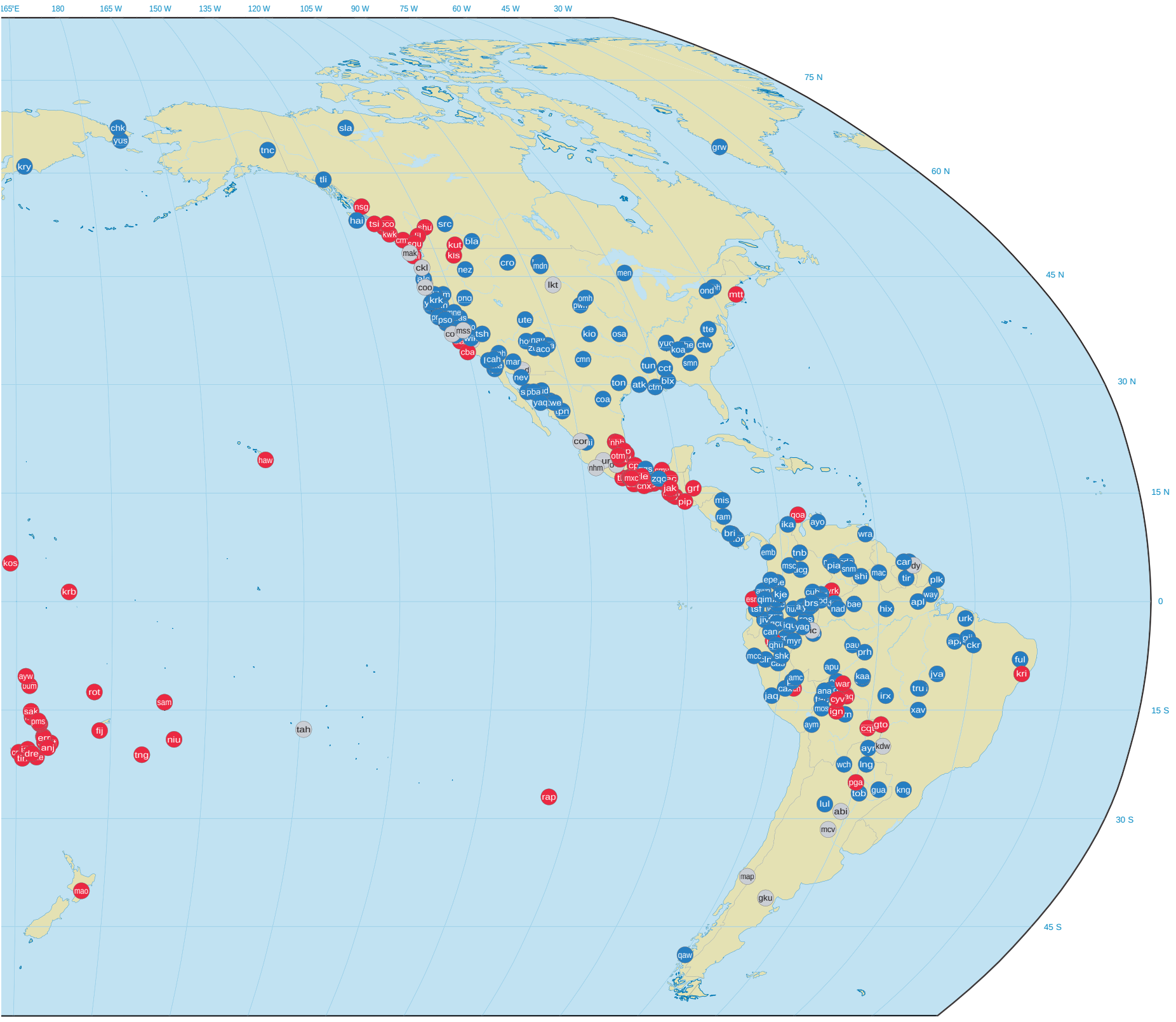
86 Order of Genitive and Noun

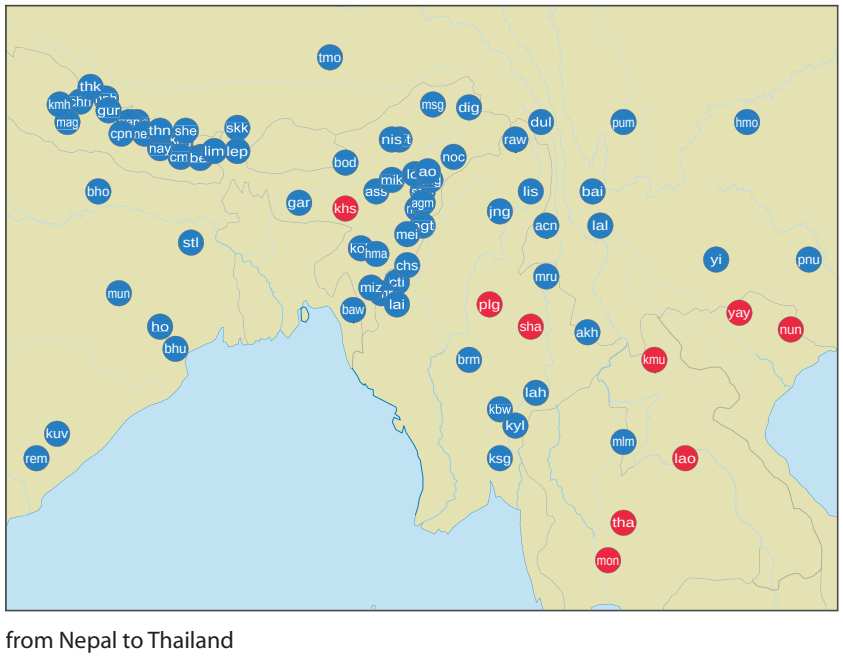
MATTHEW S. DRYER

Genitive Noun [608]
Noun–Genitive [415]
No dominant order [82]

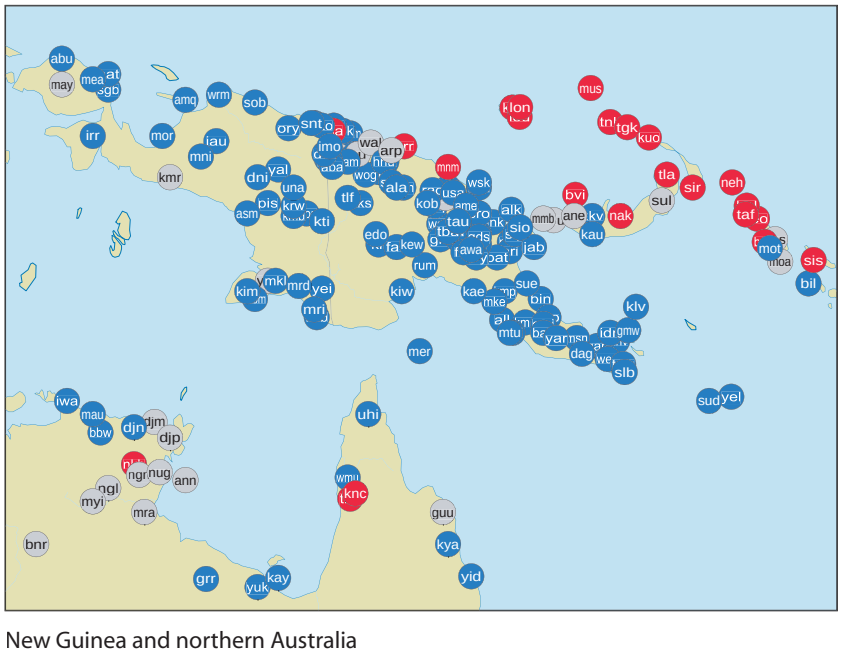
Nigeria and surrounding areas

eastern Africa





from Nepal to Thailand



New Guinea and northern Australia

87 Order of Adjective and Noun

MATTHEW S. DRYER

1 Defining the values

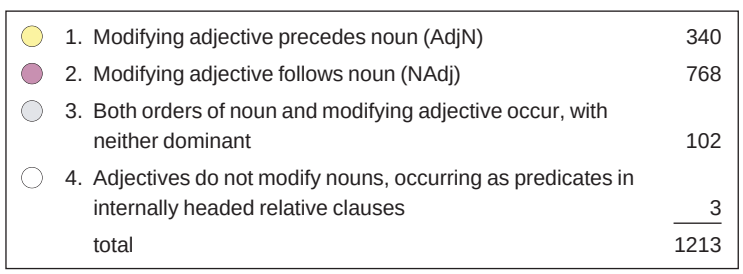
This map shows the distribution of the two possible orders of modifying adjective and noun. English is an example of a language which is **AdjN**, with the adjective preceding the noun (as in *large dogs*). Another example is Mising (Tibeto-Burman; north-east India), as in (1).

(1) Mising (Prasad 1991: 69)
azźnë *dźluŋ*
small village
Adj N
'a small village'

Examples of **NAdj** languages, with the adjective following the noun, are given in (2); Apatani, another Tibeto-Burman language spoken in north-east India, is illustrated in (2a), while Temiar (Aslian, Mon-Khmer; Malaysia) is illustrated in (2b).

(2) a. Apatani (Abraham 1985: 23)
aki *atu*
dog small
N Adj
'the small dog'

b. Temiar (Benjamin 1976: 155)
dēk *mənūʔ*
house big
'big house'



In some languages, both orders of adjective and noun occur. In some of these, an argument can be given that one of the two orders is dominant (see the box section "Determining Dominant Word Order" on p. 371). For example, in Huasteca Nahuatl (Uto-Aztecan; Mexico) the words for 'good' and 'big' precede the noun, but other adjectives more frequently follow the noun (Beller and Beller 1977: 233). This is taken here as a basis for saying that NAdj order is dominant in Huasteca Nahuatl, and it is thus shown on the map as NAdj. Tagalog (Austronesian; Philippines), in contrast, is shown on the map as an instance of a language of the third type, having **both orders with neither order dominant**, because there is no evidence from the source that one of the orders is dominant (Schachter and Otanes 1972: 118, 121–2).

It should be emphasized that this map shows the order of adjectives *modifying* a noun. It does not show the order of noun and predicative adjective, when the noun is subject and the adjective is functioning as the predicate, as in English *the boy is tall* and in the example in (3) from Simeulue (Austronesian; Sumatra, Indonesia).

(3) Simeulue (Kähler 1963: 131)
mexiao *luan* *ere*
clean river this
'This river is clean.'

The adjective *mexiao* 'clean' in (3) is not modifying the noun *luan* 'river'; rather, *luan ere* 'this river' is the subject and *mexiao* 'clean' is the predicate. The position of *mexiao* 'clean' in (3) reflects the fact that it is functioning as the predicate and predicates precede their subjects in Simeulue. Adjectives modifying nouns, in contrast, *follow* the noun in Simeulue, as illustrated by the adjective *tu'a-tu'a* 'very old' following *ata* 'person' in (4).

(4) Simeulue (Kähler 1963: 17)
ŋaŋ *sa'a* *bəsaŋ* *sara* *ata* *tu'a-tu'a*
already then come one person old-old
'Then a very old man came.'

For the purposes of this map, the term *adjective* should be interpreted in a semantic sense, as a word denoting a descriptive property, with meanings such as 'big', 'good', or 'red'. It does not include nondescriptive words that commonly modify nouns, such as demonstratives (like *this* in *this dog*) (see Map 88), numerals (as in *two dogs*) (see Map 89), or words meaning 'other' (as in *the other dog*). In some languages, like English, adjectives form a distinct word class. In other languages, however, adjectives do not form a distinct word class and are verbs or nouns (see Chapter 118). For example, in Eastern Ojibwa (Algonquian; eastern Canada and United States), words expressing adjectival meaning are just like verbs morphologically and syntactically. The example in (5a), involving a word meaning 'tall' being used predicatively, inflects for a first-person singular subject with a prefix *n-* in the same way as the inflection for the verb meaning 'sing' in (5b).

(5) Eastern Ojibwa (Rich Rhodes, p. c.)
a. *n-ginooz*
1SG-tall
'I am tall.'
b. *n-nagam*
1SG-sing
'I am singing.'

Similarly, these two words inflect in the same way when they are used attributively to modify a noun, as in (6).

(6) Eastern Ojibwa (Rich Rhodes, p. c.)
a. *nini* *e-gnoozi-d*
man REL-tall-3SG
'a tall man'
b. *nini* *e-ngamo-d*
man REL-sing-3SG
'a man who is singing'

Both modifying words in (6) bear third-person subject marking and a relativizing prefix *e-*. Because words expressing adjectival meaning are really verbs in Ojibwa, instances in which such words modify nouns, like (6a), are, strictly speaking, relative clauses. In other languages, words expressing adjectival meaning form a well-defined subclass of verbs, sharing certain grammatical properties with other verbs, but differing in other respects. For example, in Lealao Chinantec (Oto-Manguean; Mexico) there is a subclass of verbs that express adjectival meanings which occur with verbal inflections but which differ from other verbs in that they can directly modify nouns without a relative marker. Compare (7a), with a nonadjectival verb preceded by the relative marker *ʔi*^{M}, with (7b), in which an adjectival verb immediately follows the noun without the relative marker.

(7) Lealao Chinantec (Rupp 1989: 86)
a. *mïʔ*^{M} *[ʔi*^{M} *ka*^{L}*-lá*^{H} *mi*^{VH}*-liuʔ*^{H}*]*
clothes REL PST-buy.3 CLF-little
'the clothes that the child bought'
b. *mï*^{VH}*-kuï:*^{M} *tia:*^{M}
CLF-corn white
'white corn'

For the purposes of this map, these distinctions in word class are ignored: a word is treated as an adjective, regardless of its word class

in the language, as long as it denotes a descriptive property. The map also ignores the question of whether the adjectives are modifying nouns directly or whether they are the predicate of a relative clause which is modifying the noun. It is a matter for future research to determine whether any of these distinctions provide a basis for further patterns in the distribution of AdjN and NAdj order, either typologically or geographically.

The fourth type shown on the map are languages in which the **adjectives do not modify nouns**, in which in the closest equivalent to such structures, the adjective is actually the predicate in an internally headed relative clause (see Chapter 90), and the noun is serving as its subject. Internally headed relative clauses in Mesa Grande Diegueño (Yuman; southern California and north-west Mexico) are illustrated in (3) in Chapter 90. The example in (8a) below illustrates the translation equivalent of an adjective modifying a noun, but in fact the word for 'white' in (8a) is the verbal predicate of an internally headed relative clause and the word *aq* 'bone' is functioning as the subject of that verb; its structure is exactly parallel to the structure in (8b), with a nonadjectival intransitive verb.

(8) Mesa Grande Diegueño (Couro and Langdon 1975: 224, 236)

a. *'iikwich=ve=ch* *[aq* *ku-nemshap]=vu* *aakwal*
man=DEF=SUBJ [bone REL.SUBJ-white]=DEF lick
'The man licked the white bone.'

b. *kwenychekwii=ve=ch* *[hekwany* *ku-mii]=vu*
old.woman=DEF=SUBJ [baby REL.SUBJ-cry]=DEF
selyewelyuu
tickle
'The old woman tickled the baby that cried.'

While superficially it might not be obvious that the examples in (8) involve internally headed relative clauses, the fact that they have exactly the same form as the examples in (3) in Chapter 90, which are clearly internally headed, means that these examples apparently involve internally headed relative clauses as well. Note that the order of the noun and adjective in (8a) simply reflects the normal order of subject and verb in Diegueño. Languages in which adjectives do not really modify nouns, but are predicates in internally headed relative clauses, are probably more common than the map suggests, both because grammarians have until recently often failed to recognize internally headed relative clauses, and because the simple structure of internally headed relative clauses with just noun plus adjective is such that it may not be recognized that they are simple instances of internally headed relative clauses. Some of the languages that are shown as AdjN or as NAdj may prove under more careful analysis to be better treated as languages in which the adjectives are predicates in internally headed relative clauses.

2 Geographical distribution

Both AdjN and NAdj orders are common in the world, though there are more than twice as many NAdj languages on the map. There are also clear geographical patterns. NAdj order is overwhelmingly the dominant order in Africa, though there exist a few well-defined pockets of AdjN order. This area of NAdj order in Africa can be seen as extending northward into south-west Europe and to the north-east into the Middle East. NAdj order is also the dominant type in a large region stretching from north-east India through South-East Asia eastward among Austronesian languages into the Pacific, except in the Philippines. It is the dominant order in both New Guinea and Australia, though there are many exceptions. Both orders are common in North America, but NAdj order is noticeably more common in the eastern half of the United States and among the more centrally located languages of Mesoamerica. NAdj is the majority type in South America, again with many scattered exceptions.

By far the largest area in which AdjN is found is a large area covering much of Europe and Asia, except in south-west Europe, the Middle East, and South-East Asia. Within this area, there are relatively few exceptions to the dominance of AdjN order, though a number of Tibeto-Burman languages of the Himalayan region are NAdj. AdjN order is clearly a minority type in Africa, but there are some clear pockets, notably in Ethiopia, in central Africa, and among Khoisan languages in southern Africa. Similarly, AdjN is a minority type in Australia, though there is a scattering of them, including pockets in the south-east and in the middle of the north coast. The situation is similar in New Guinea, with a couple of pockets of AdjN order in the eastern Highlands and in the lower Sepik valley. AdjN order is as common as NAdj order in North America, and is more common in the western regions of Canada and the United States. Both orders are found in Mesoamerica, though AdjN order is more common in the south-east. In South America, the AdjN languages are primarily confined to the western half of the continent.

While some of the geographical patterns shown on the map reflect areal phenomena that cross genealogical boundaries, there are cases in which knowing genealogical classification can explain instances where languages in the same area are of different types. For example, Romanian is a NAdj language surrounded by AdjN languages, but this reflects the fact that it is a Romance language separated from other Romance languages and like other Romance languages is NAdj.

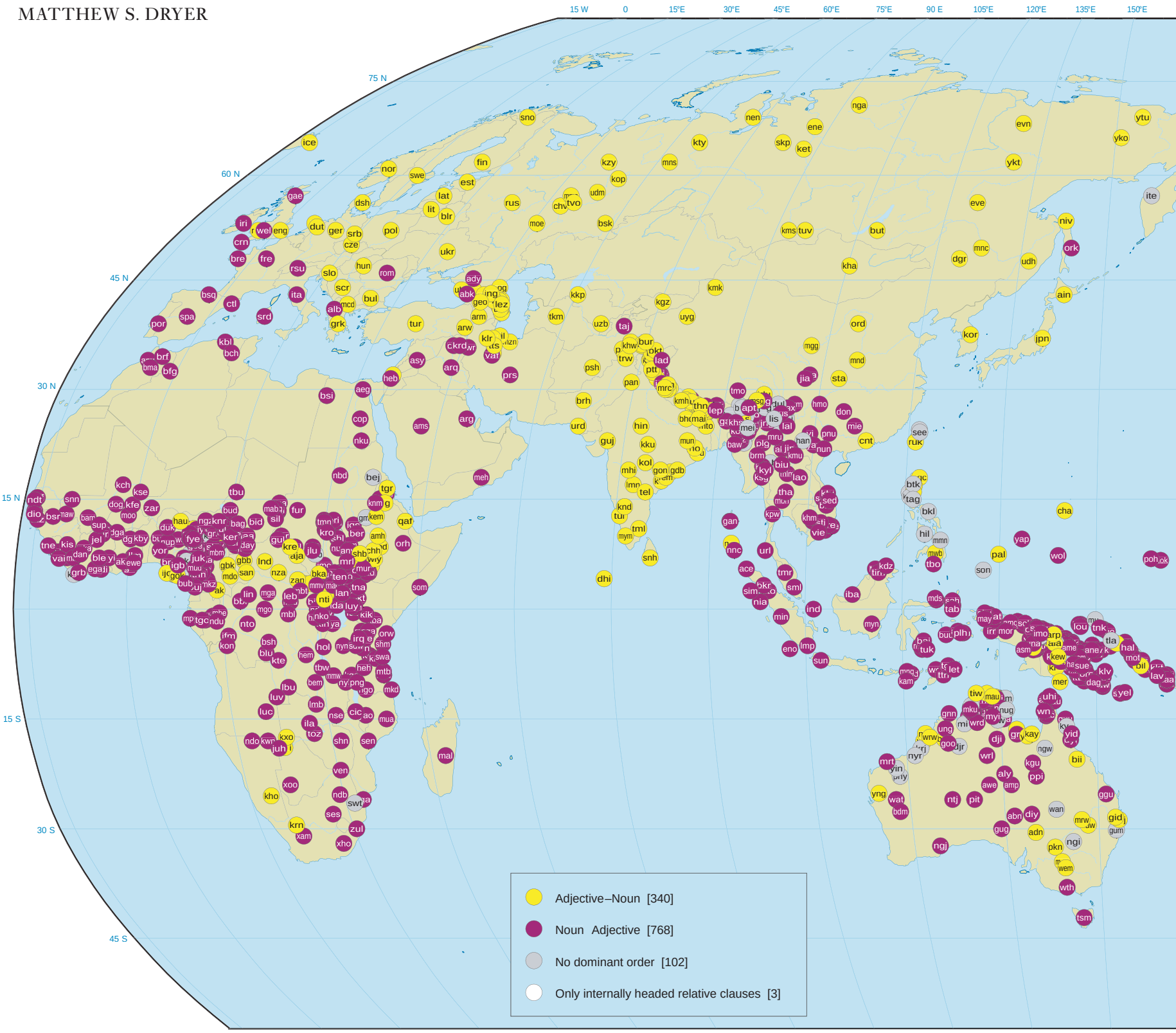
Languages lacking a dominant order of adjective and noun are widely scattered, but are noticeably more common in the Philippines, in an area in and around Myanmar, in Australia, and in the Americas. The languages shown as lacking constructions with an adjective modifying the noun because the closest equivalent involves an internally headed relative clause are all in the Americas.

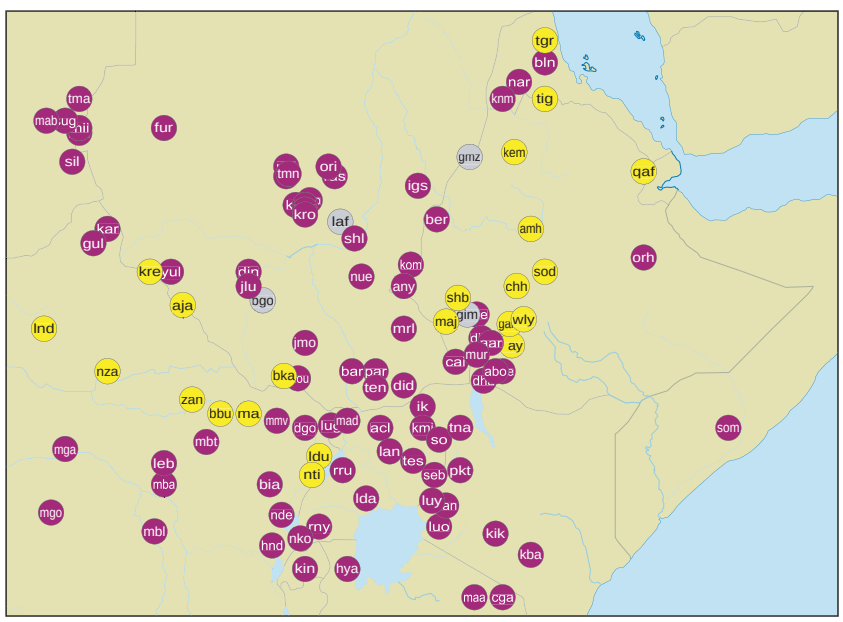
3 Theoretical issues

The order of adjective and noun has been of most interest because it is often thought that it correlates with the order of object and verb. However, as shown by Dryer (1988a, 1992) and in Chapter 97 of this atlas, this is not the case: NAdj order is more common than AdjN order, both among OV languages and among VO languages. Greenberg (1963) and Hawkins (1983) discuss other cross-linguistic generalizations involving the order of adjective and noun. As noted above, it remains to be investigated whether distinguishing among different sorts of languages on the basis of the extent to which adjectives are a distinct word class (or subclass) might lead to new generalizations relating to the order of adjective and noun.

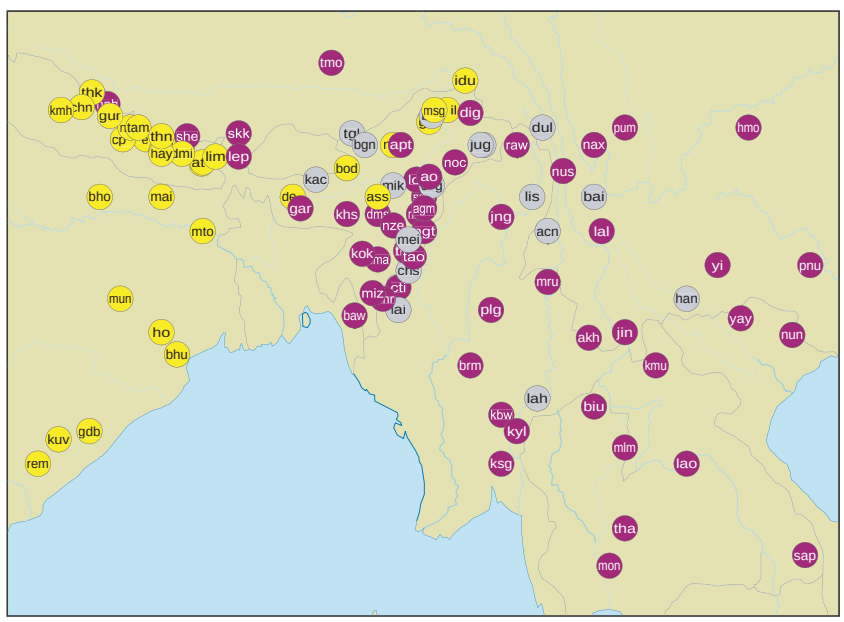
87 Order of Adjective and Noun

MATTHEW S. DRYER

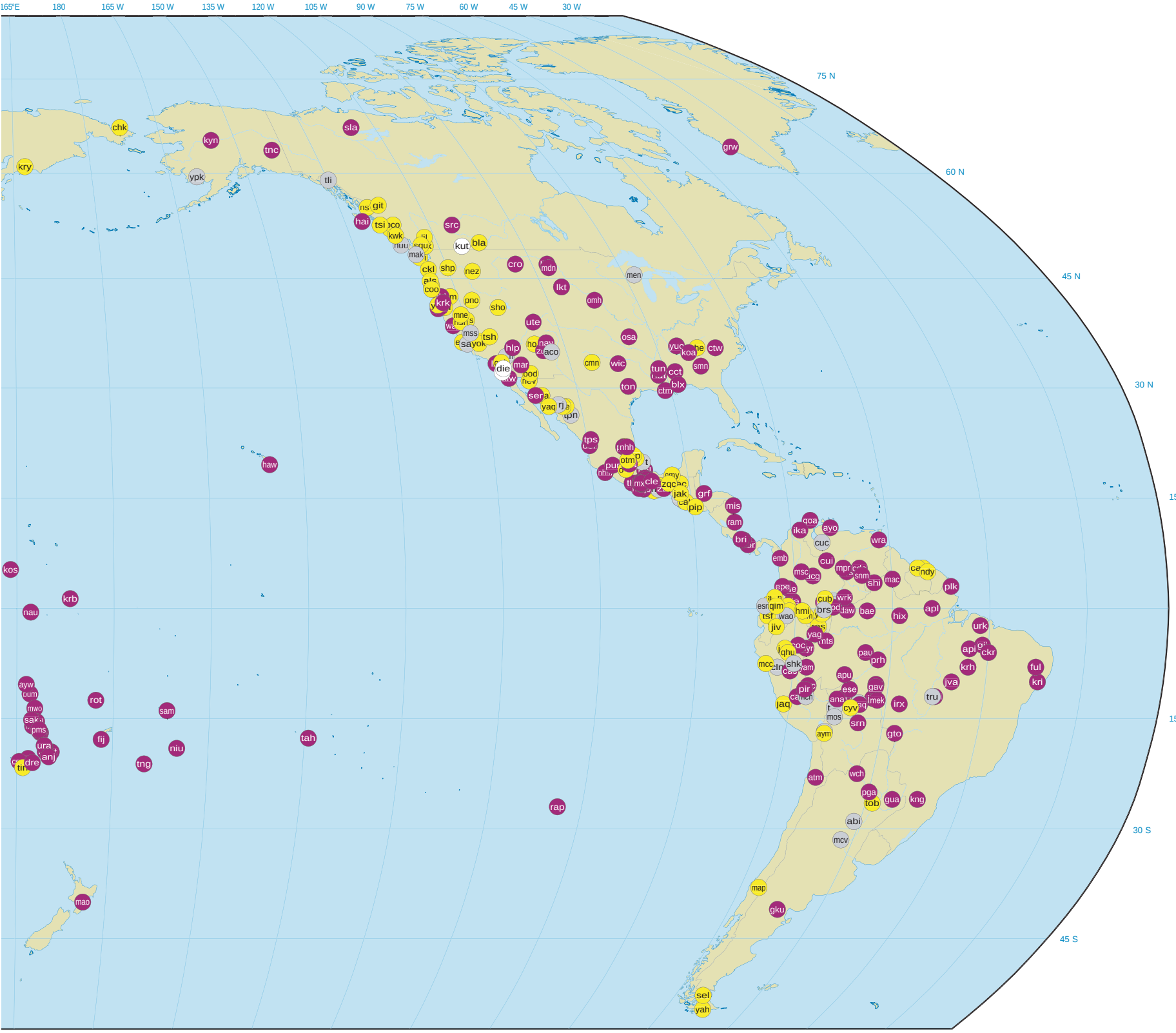


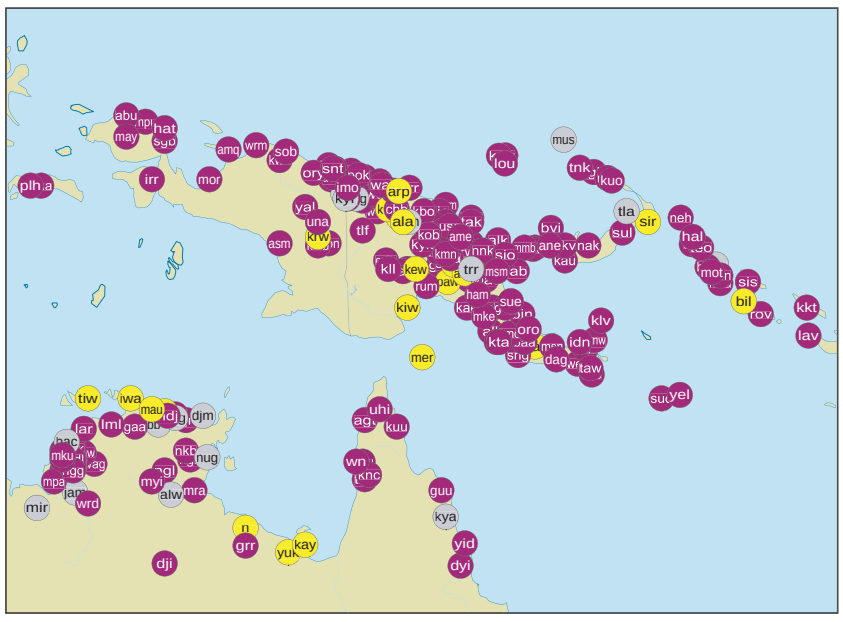


eastern Africa

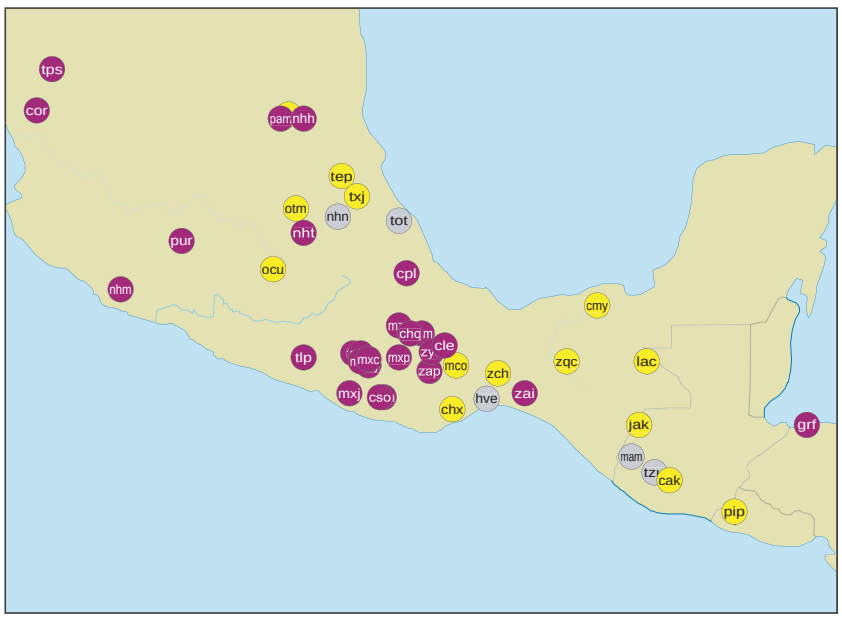


from Nepal to Thailand





New Guinea and northern Australia



Mesoamerica

88 Order of Demonstrative and Noun

MATTHEW S. DRYER

1 Defining the values

This map shows the different orders of demonstrative and noun. For the purposes of this map, a word or affix is considered a demonstrative if it satisfies at least one of two criteria: (1) it has among its uses a meaning that contrasts with some other form in terms of physical proximity to the speaker, so that there is at least a two-way contrast of proximal (near speaker) versus distal (not near speaker); or (2) the form has among its uses an indication that the hearer is intended to direct their attention towards something in the physical environment. In many and probably most languages, demonstratives are associated with both of these functions. The words *this* and *that* (and their plural forms *these* and *those*) constitute the demonstratives of English. In many languages, including English, the demonstratives can be used without a difference in form, either as modifiers of nouns, as in (1a), or pronominally (not modifying a noun), as in (1b).

(1) a. *I want that book.*
b. *I want that.*

This map shows the order of such words with respect to the noun when they accompany a noun, as in (1a). In some languages, the form of demonstratives depends on whether they are being used pronominally or with a noun; see Map 42 for more on this phenomenon.

1. Demonstrative word precedes noun (DemN)	496
2. Demonstrative word follows noun (NDem)	478
3. Demonstrative prefix on noun	9
4. Demonstrative suffix on noun	29
5. Demonstrative simultaneously before and after noun	15
6. Two or more of above types with none dominant	58
total	1085

The first type shown on the map consists of languages in which **the demonstrative is a separate word which precedes the noun**. English is an instance of such a language, as in (1a), as is Fore (East New Guinea Highlands, Trans-New Guinea; Papua New Guinea), as in (2).

(2) Fore (Scott 1978: 89)
má: *ntamá*
this house
Dem N
'this house'

The second type includes languages in which **the demonstrative is a separate word which follows the noun**. An example of a language of this type is Maba (Maban, Nilo-Saharan; Chad), as in (3).

(3) Maba (Trenga 1947: 70)
mašuk *wak* *kulak* *ti*
man this tall be.3SG
N Dem
'This man is tall.'

The third type consists of languages in which **the demonstrative is a prefix** on the noun. An example of such a language is Abzakh Adyghe (North-West Caucasian; Caucasus region, Russia), illustrated by the prefix *mə-* 'this' in (4).

(4) Abzakh Adyghe (Paris 1989: 175)
mə-sɔy-ǧʷəneǧʷə-r
this-1SG.POSS-neighbour-DEF
'this neighbour of mine'

The fourth type consists of languages in which **the demonstrative is a suffix on the noun**, as in (5) from Gude (Chadic; Nigeria).

(5) Gude (Hoskison 1983: 45)
a. *zəmə-na* b. *zəmə-ta*
food-this food-that.far
'this food' 'that food (far)'

The fifth type involves languages where the expression of demonstrative meaning involves **a demonstrative word or affix preceding the noun occurring simultaneously with a demonstrative word or affix following the noun**. An example of such a language is Milang (Tibeto-Burman; north-east India), as illustrated in (6).

(6) Milang (Tayeng 1976: iv)
yo *miu* *yo*
this boy this
'this boy'

In Nishi (Tibeto-Burman; north-east India), the demonstratives that precede are different from those that follow the noun, as in (7); the ones that precede are the words that otherwise function as demonstrative adverbs (*here, there*).

(7) Nishi (Hamilton 1900: 20)
sa *mindui* *sî*
here buffalo this
'this buffalo'

In Lai (Tibeto-Burman; Myanmar), there is an invariant demonstrative word that precedes the noun, while the demonstratives that follow vary for distance, as in (8).

(8) Lai (Hay-Neave 1953: 37, 44)
a. *mah* *lam* *hi*
DEM road this
'this road'
b. *mah* *mipa* *khi*
DEM man that
'that man'

In Chai (Surmic, Nilo-Saharan; Ethiopia), the demonstrative takes the form of a simultaneous prefix and suffix. The prefix is constant, but there are two suffixes, one with proximal meaning, the other with distal meaning, as illustrated in (9).

(9) Chai (Last and Lucassen 1998: 400)
a. *ŋà-sábbá-à*
DEM-head-this
'this head'
b. *ŋà-ɓíré-nù*
DEM-stick-that
'that stick'

In some languages, the pattern of demonstratives simultaneously preceding and following the noun is optional. An example of such a language is Bawm (Tibeto-Burman; India and Bangladesh); in (10a), there are demonstratives both preceding and following the noun, while in (10b), there is only a postnominal demonstrative.

(10) Bawm (Reichle 1981: 136, 137)
a. *hi* *Pathian* *biakin* *hi*
this God temple this
'this temple of God'
b. *khuavâng* *kha*
prophet that
'that prophet'

Such languages are shown on the map according to the position in which the demonstrative is obligatory; hence Bawm is shown as placing the demonstrative after the noun.

The sixth type includes languages in which **two or more of the above constructions occur without either being dominant**. Most of these languages are ones in which both orders of

demonstrative and noun are common, as in Gulf Arabic, illustrated in (11).

(11) Gulf Arabic (Holes 1990: 175)

a. *il-qaraar* *haadha*
the-decision this
'this decision'

b. *haadha* *l-qaraar*
this the-decision
'this decision'

Tigre (Semitic; Eritrea) represents a fairly rare subtype of the type in which neither order is dominant, in that the proximal and distal demonstratives occur in different positions: the proximal one obligatorily precedes the noun while the distal one normally follows, as in (12).

(12) Tigre (Raz 1983: 45)

a. *ʔəllan* *ʔamʕəlāt*
this.F.PL days
'these days'

b. *ʔəb* *laʔawkād* *lahay*
at time that M
'at that time'

Also treated as languages in which no one type is dominant are languages with two constructions, one involving an affix, the other involving a separate word. For example, Doyayo (Adamawa, Niger-Congo; Cameroon) has both demonstrative suffixes, as in (13a), and separate demonstrative words that occur at the end of noun phrases, as in (13b).

(13) Doyayo (Wiering 1994: 148, 170)

a. *i³s-ɔ²³*
goat-that
'that goat'

b. *ɯa²lɛ³* *sɛ¹rɛ³* *rɔ⁴bɛ¹* *gbɔ¹lɛ³* *ee¹rɛ¹* *ɯɔ¹rɔ¹*
man.PL young foreign big two those
'those two big young aliens'

The demonstrative suffix and the demonstrative word in Doyayo cannot co-occur in the same noun phrase. It is not clear what conditions the choice between them.

The order of demonstrative and noun is more often fixed than is the case with many other pairs of elements. In many languages in which word order is in general flexible within the noun phrase, the order of demonstrative and noun is fixed. For example, in Mangarrayi (Northern Territory, Australia), adjectives, numerals, and genitives can either precede or follow the noun, but the demonstrative must precede the noun, as in (14).

(14) Mangarrayi (Merlan 1982: 63)
ŋaḷi-na *ŋala-gaḍugu*
F.NOM-that F.NOM-woman
'that woman'

While this map shows the order of demonstrative and noun and excludes words that are specifically definite articles, there are many languages in which one of the demonstratives, most commonly the distal demonstrative, is very frequently used anaphorically in ways that resemble a definite article. See Chapter 37. However, such a word or affix will be treated as a demonstrative as long as it satisfies one of the two criteria described in the first paragraph of this chapter as definitional for demonstratives. It should be noted that languages sometimes have words that are called *demonstratives* in grammars but which lack both of the two functions treated as definitional here; this is generally because they belong to a class of words that also includes words that do satisfy the definition assumed here. For example, in Mauka (Mande; Côte d'Ivoire; Ebermann 1986: 76), there is a word that Ebermann calls a demonstrative pronoun that occurs either as a pronoun or as a modifier of the noun, but whose meaning is necessarily anaphoric in the discourse context, meaning 'the aforementioned'. Since it lacks a contrast involving proximity and since it is not used to direct the hearer's attention to something in the physical environment, I do not treat it as a demonstrative.

2 Geographical distribution

Languages with demonstrative words preceding the noun are the overwhelmingly dominant type in most of Europe and Asia, except in South-East Asia and except for a few languages in western Europe (Celtic languages and Basque). This is also the dominant type in the Americas, though there are many scattered exceptions. Languages with demonstrative words following the noun are the overwhelmingly dominant type in Africa and in a large area stretching from South-East Asia eastward into the Pacific. Both of these types are common in Australia and in New Guinea, and there are complex areal patterns within Australia. Languages with demonstrative affixes are not common, though they are somewhat more common in Africa, and languages with prefixes are more common in areas where demonstrative words tend to precede the noun while languages with suffixes are more common in areas where demonstrative words tend to follow the noun, as in Africa. Languages with demonstratives preceding and following the noun are not common; they are found in widely scattered areas, with a notable pocket among Tibeto-Burman languages in north-east India. Languages in which none of the preceding types is dominant are widely scattered, somewhat more common than average in Africa and somewhat less common in Europe and Asia.

3 Theoretical issues

The term *determiner* is often used, especially with reference to English, for a class of words that includes demonstratives but often other words as well, such as articles (words like *the* and *a* in English). The concept of determiner is well motivated for English, since there is a single syntactic position in the noun phrase which can be filled by at most one determiner. But for many other languages, there is less motivation for a grammatical class of determiners, since the language may lack articles, or it may have articles but they occur in a different position in the noun phrase from demonstratives (see Chapters 37 and 38 on articles). For example, in Kana (Niger-Congo; Nigeria), the definite article precedes the noun, while the demonstrative follows, and the two can co-occur, as in (15).

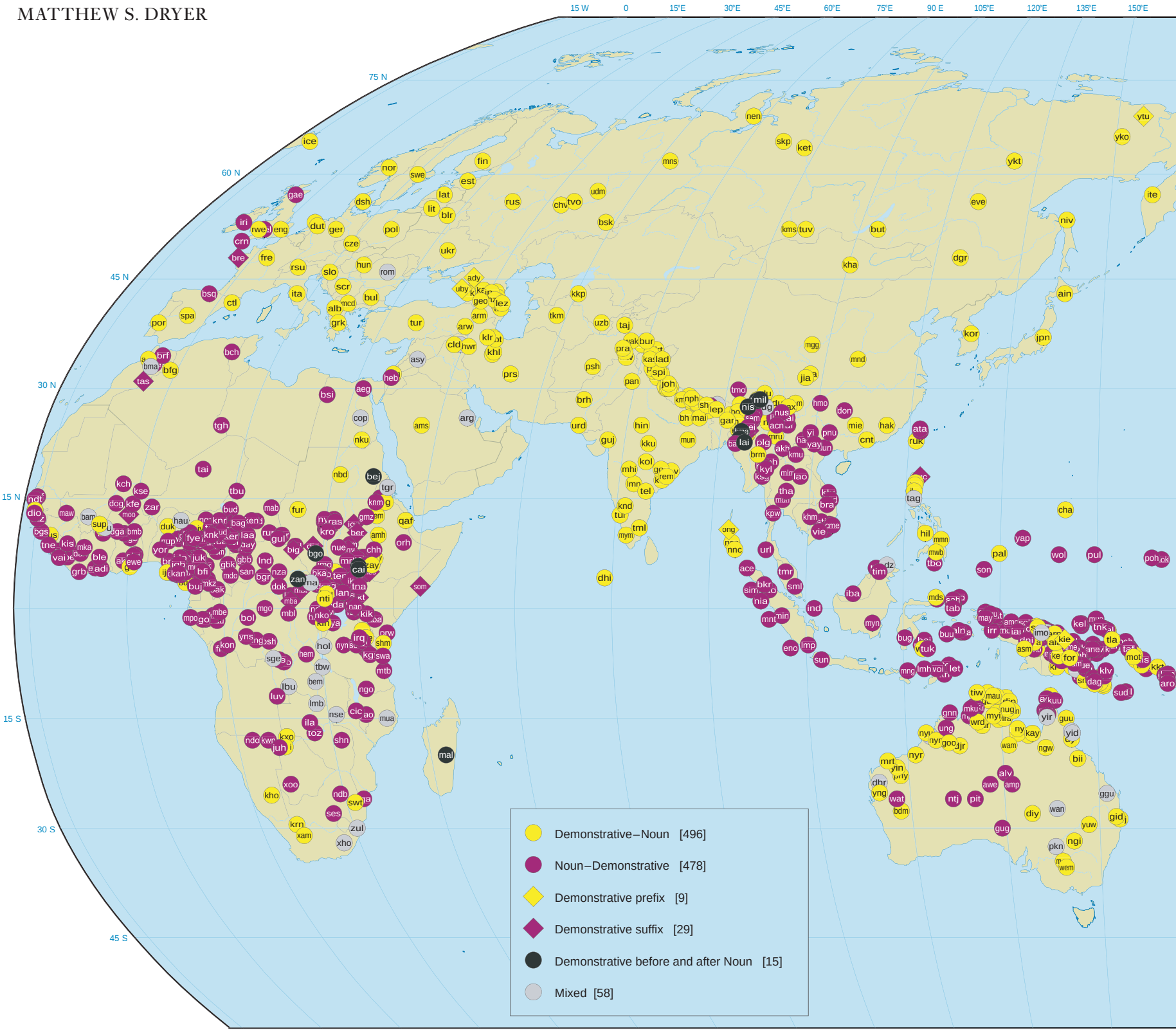
(15) Kana (Ikoro 1996: 70)
ló *bárí* *āmā*
DEF fish this
'this fish'

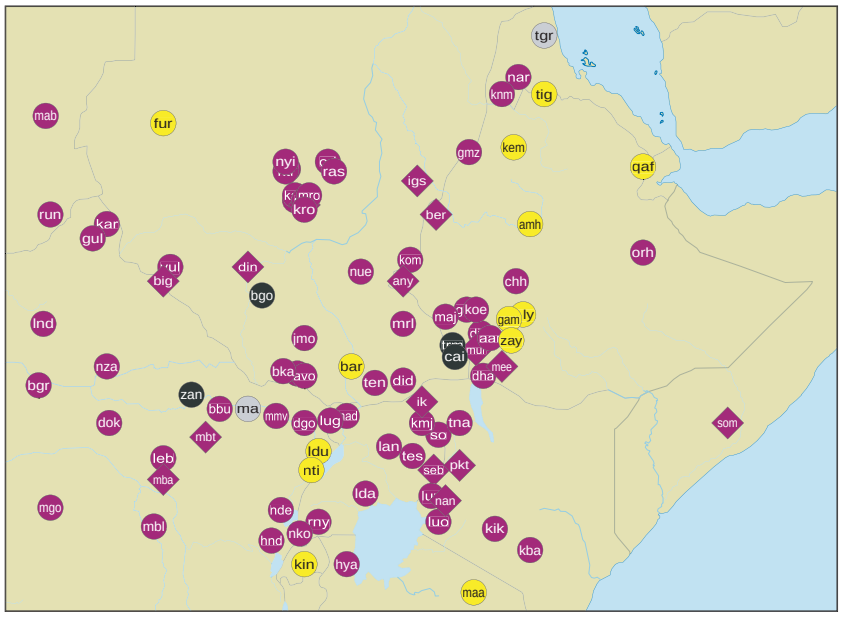
Similarly, in Ngiti (Central Sudanic; Democratic Republic of Congo), the demonstrative and definite article both precede the noun, but can co-occur, as in (16).

(16) Ngiti (Kutsch Lojenga 1994: 376)
yà *nd̶ɨ* *dza*
this DEF house
'this house (mentioned before)'

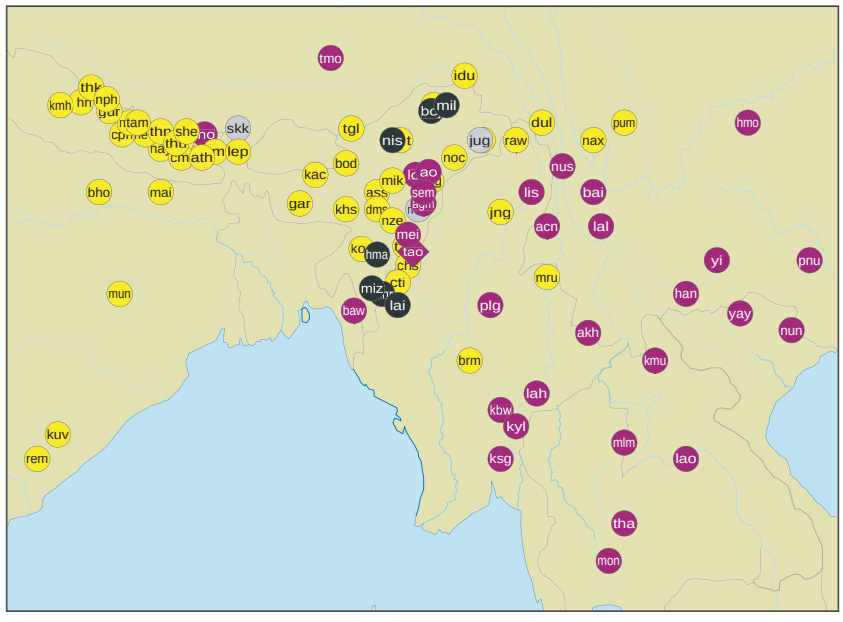
88 Order of Demonstrative and Noun

MATTHEW S. DRYER

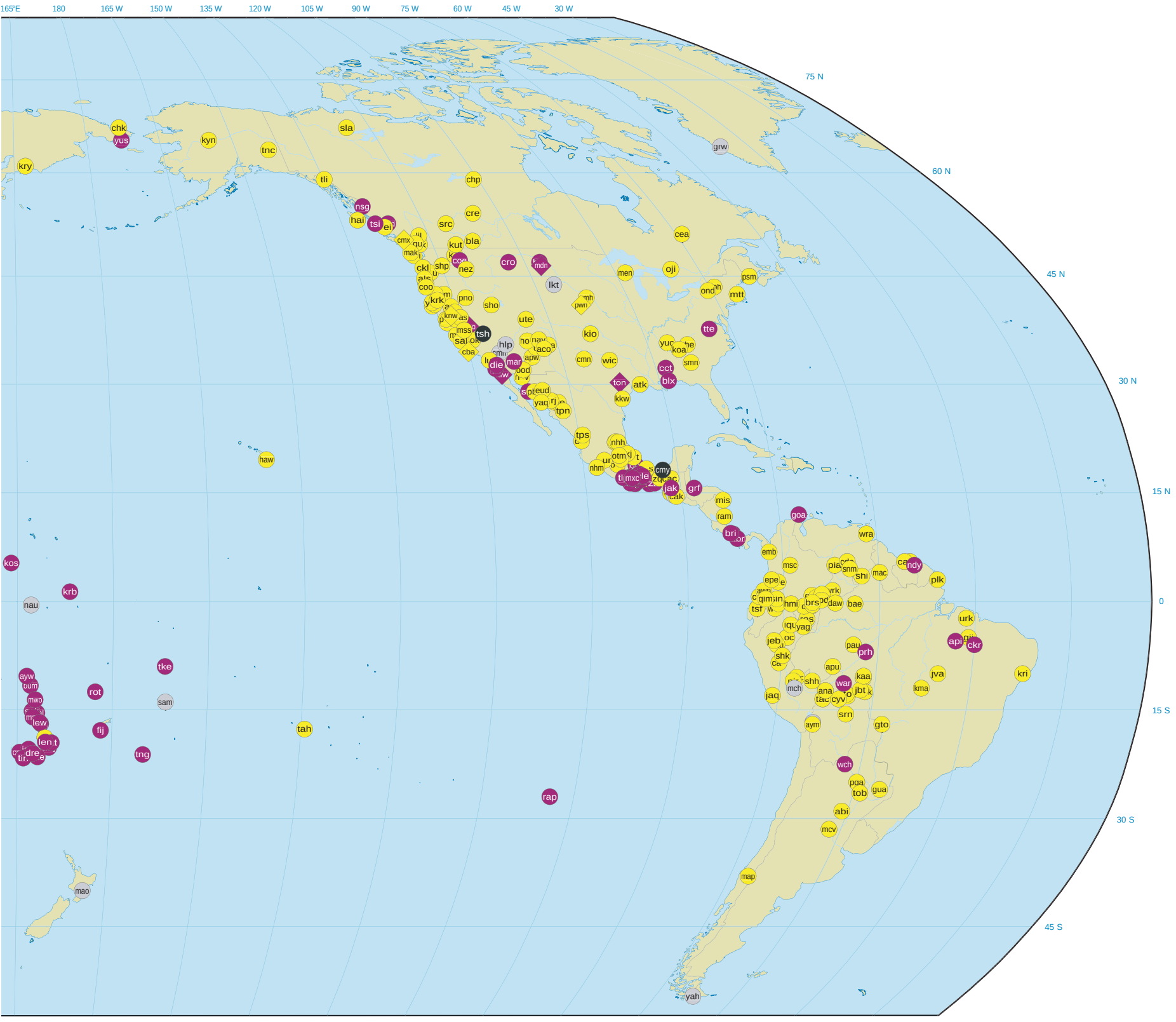


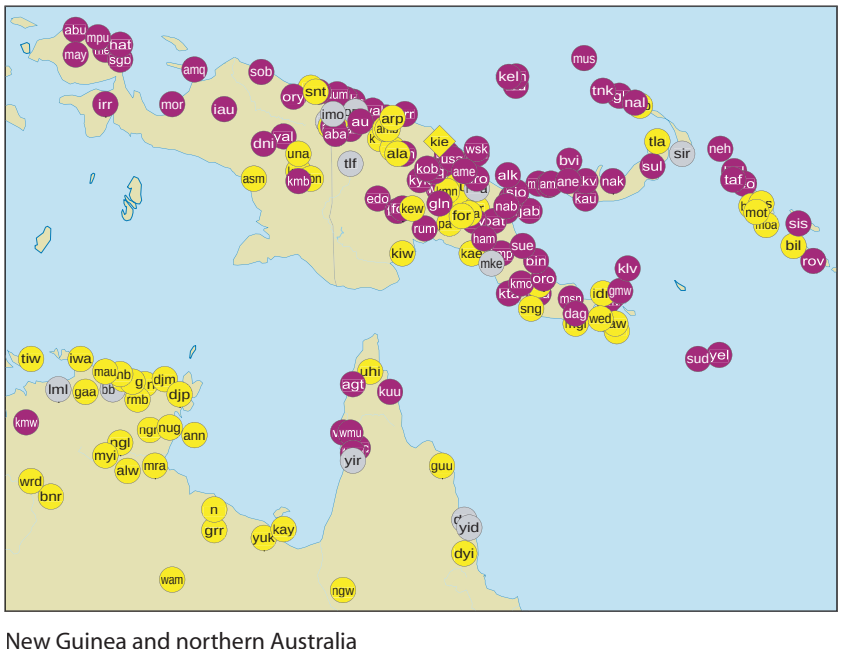


eastern Africa

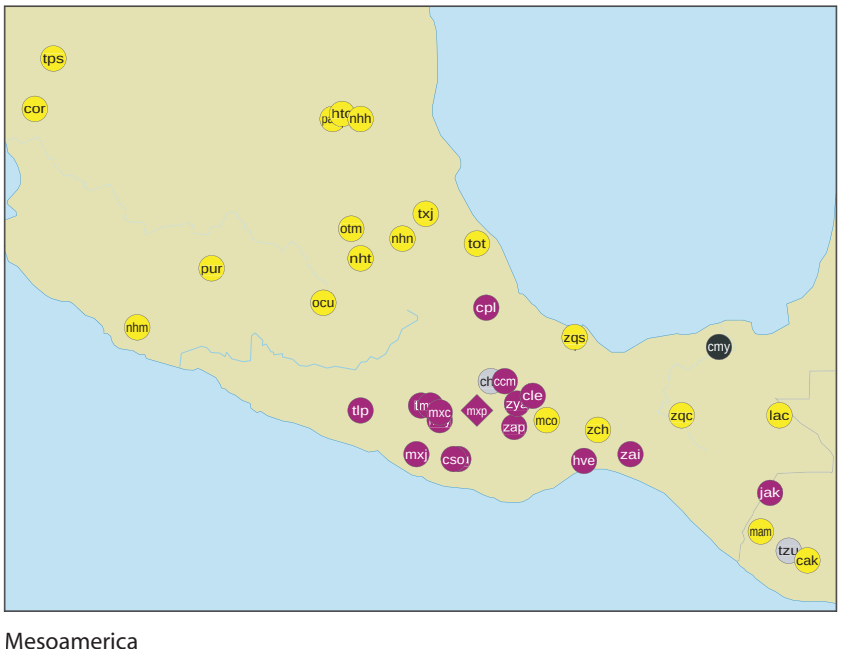


from Nepal to Thailand





New Guinea and northern Australia



Mesoamerica

89 Order of Numeral and Noun

MATTHEW S. DRYER

1 Defining the values

This map shows the order of cardinal numerals with respect to a noun they modify. Cardinal numerals are words denoting the number of things referred to, as in English *four chairs*, in contrast to ordinal numerals, which specify the rank in some order of objects, as in English *the fourth chair*. See also Map 131 on numeral bases and Map 53 on ordinal numerals.

1.	Numeral precedes noun (NumN)	430
2.	Numeral follows noun (NNum)	515
3.	Both orders of numeral and noun with neither order dominant	54
4.	Numeral only modifies verb	2
	total	1001

The first type shown on the map consists of languages in which **the numeral precedes the noun**. English is an instance of such a language, as is Selknam (Chon; Tierra del Fuego, Argentina), illustrated in (1).

(1) Selknam (Najlis 1973: 28)

şèàwken-èn	*naʔ*
three-F	woman
Num	N

'three women'

The second type consists of languages in which **the numeral follows the noun**. Examples of such languages include Pumi (Tibeto-Burman; China), as in (2a), and Ju|'hoan (Northern Khoisan; Namibia), as in (2b).

(2) a. Pumi (Ding 1998: 191)

qüa	*xüé*
pig	eight
N	Num

'eight pigs'

b. Ju|'hoan (Dickens 1992: 61)

jú	*tsán*	*gèà*	*tjù*	*n!áng*
person	two	be	house	in
N	Num			

'There are two people in the house.'

The third type shown on the map are languages in which **both orders of numeral and noun occur with neither order dominant** (see the box section "Determining Dominant Word Order" on p. 371). An example of such a language is Acoma (Keresan; New Mexico), as in (3).

(3) Acoma (Maring 1967: 119)

a.

tʸánakʼa	*wáaštíṭa*
four	fawn

'four fawns'

b.

kákʰana	*tʰúwé*
wolf	two

'two wolves'

In most languages with both orders of numeral and noun, it is not clear whether there is any difference in meaning associated with the two orders. In Nias (Austronesian; Sumatra, Indonesia), the noun phrase is interpreted as indefinite when the numeral precedes the noun, as in (4a), but definite when it follows, as in (4b).

(4) Nias (Brown, forthcoming)

a.

öfa	*geu*	*m-baβi*	*s=afusi*
four	CLF	ABS-pig	REL=white

'four white pigs'

b.

baβi-ra	*s=afusi*	*si=öfa*	*geu*
pig-3PL.POSS	REL=white	REL=four	CLF

'their four white pigs'

In some languages there is a split among numerals, some preceding the noun, others following. For example, in Egyptian Arabic, the numerals for 'one' and 'two' follow the noun, while higher numerals precede the noun, as in (5).

(5) Egyptian Arabic (Gary and Gamal-Eldin 1982: 111)

a.

binteen	*ʔitneen*
girl.DU	two

'two girls'

b.

talat	*banaat*
three	girl.PL

'three girls'

This variety of Arabic is shown on the map as lacking a dominant order of numeral and noun. In Moroccan Arabic, only the numeral for 'one' follows the noun, as in (6a), while all other numerals precede the noun, as in (6b).

(6) Moroccan Arabic (Harrell 1962: 206)

a.

mṛa	*weḥda*
woman	one.F

'one woman'

b.

žuž	*ktub*
two	book

'two books'

When only one numeral differs from the others in position, as in Moroccan Arabic, the position of the other numerals is treated as dominant. Hence, Moroccan Arabic is shown on the map as having numeral before noun as the dominant order. Iraqi Arabic is intermediate between Egyptian and Moroccan Arabic: the numeral for 'one' follows, the numeral for 'two' either precedes or follows, and numerals higher than two normally precede (Erwin 1963: 348). It is shown on the map as lacking a dominant order of numeral and noun.

In most languages in which some numerals precede the noun while others follow, both of the two sets of numerals represent continuous sequences of numbers, so that there is some number *n* such that numerals from 1 to *n* occur on one side of the noun while numerals greater than *n* occur on the other side. For Egyptian Arabic, illustrated above in (5), the value of *n* is two. For Tauya (Madang, Trans-New Guinea; Papua New Guinea), the value of *n* is four (MacDonald 1990: 117–18). For Nivkh (isolate; Sakhalin Island, Russia), the value of *n* is five (Panfilov 1962: 129–30). For Runga (Maban; Chad), the value of *n* is six (Nougayrol 1990: 40–2).

There are other languages which do not exhibit a pattern with a single cut-off point. The majority of exceptions appear to be languages in which the numeral for 'two' occurs on one side of the noun, while the numeral for 'one' and higher numerals occur on the other side. Examples of languages of this sort include Shipibo-Konibo (Panoan; Peru; Pilar Valenzuela, p.c.) and Coptic (Egyptian; Lambdin 1983: 56). A more unusual case is provided by Obolo (Benue-Congo; Nigeria), in which the numerals for ten, twenty, and four hundred precede the noun, while other numerals consisting of one word follow. The special treatment of twenty and four hundred reflects the vigesimal system (see Map 131). With numerals consisting of more than one word, the first word (which will normally be one of the words for 'ten', 'twenty', or 'four hundred') will precede the noun, while the remainder will follow the noun, as in (7c).

(7) Obolo (Faraclas 1984: 27, 37)

a.

óbóp	*úwù*
four.hundred	house

'four hundred houses'

b.

úwù	*ílé*	*ítá*
house	big	three

'three big houses'

c.

étîp	*úwù*	*mè*	*gò*
twenty	house	and	five

'twenty-five houses'

Welsh is similar to this in that numerals consisting of more than one word place the first word in the numeral phrase before the noun, and the remainder following the noun (Williams 1980: 41). Because

a numeral consisting of a single word precedes the noun, and these include the most frequent numerals, Welsh is shown on the map as NumN; for analogous reasons, Obolo is coded as NNum.

The final type shown on the map is languages in which the numeral does not combine syntactically with the noun at all, but occurs **only as a modifier of the verb**. There are only two languages in my data in which this construction is the only normal way to express numerals, in close proximity to each other in western Brazil, namely Arára Karó (Tupian) and Wari' (Chapacura-Wanhan; Everett and Kern 1997: 154–5, 221, 347), the former illustrated in (8).

(8) Arára Karó (Gabas 1999: 172)

maʔwɨt	*ip*	*ʔɨy*	*matet*	*cagárokōm=nem*
man	fish	catch.IND	yesterday	two=ADVZ

'The man caught two fish yesterday.'

This construction is somewhat similar to the English construction illustrated in (9), in which the word *all* appears in a position associated with adverbial elements.

(9) *The guests have all left*.

There are many other languages that allow this option alongside a construction in which the numeral occurs in the noun phrase. For example, in Kutenai (isolate; western North America), there are two sorts of numeral words, one illustrated in (10a), in which the numeral occurs in the noun phrase with the noun (*ki=ʔas niu̯tahaɬ* 'two men'), the other illustrated in (10b), in which the numeral is a preverb modifying the verb, somewhat analogous to the Arára Karó construction in (8).

(10) Kutenai (own data)

a. *xa-s* *ki=ʔas* *niu̯tahaɬ* *ʔat*
then-OBV SUBORD=two young.man HAB
mitx-ni *niʔ-s* *numa-s*
shoot.at-INDIC the-OBV thunder-OBV
'Then two young men would shoot at the thunder.'

b. *n=ʔas-iɬ* *ʔupiɬ-iɬ-ni* *ʔa·qɬu̯maknik̓*
IND=two-PREVERB kill-PASS-IND Indian
'Two Indians were killed.'

Languages like Kutenai which have both constructions are coded according to the construction in which the numeral combines with the noun; hence Kutenai is shown as placing the numeral before the noun.

2 Geographical distribution

The two orders of numeral and noun show very clear geographical patterns. There is a large area in which almost all languages place the numeral before the noun, which covers most of Europe and Asia except for an area in South-East Asia described below, extending southeastward into Indonesia, the Philippines, and western Micronesia, and extending southwestward into the Middle East and across North Africa. It is also the dominant order in the Americas, especially in the westernmost regions of the United States and Mexico.

The order with the numeral after the noun is the overwhelmingly preferred order in sub-Saharan Africa, and the majority of exceptions are geographically peripheral within this region, namely in the far south of Africa, in the extreme north-west (in the vicinity of Senegal), and in the northeastern region of Ethiopia and Eritrea. This order also predominates in an area within South-East Asia, but the boundaries of this area are different from the boundaries of the South-East Asian area found with other word order characteristics. In particular, it excludes the languages of Vietnam and languages in southern China to the north and north-east of Vietnam and it excludes the languages of Indonesia and the Philippines. Conversely, it extends further to the north-west than some other word-order characteristics do, including Tibeto-Burman languages of north-east India, Myanmar, and adjacent areas of China. Even among the more western Tibeto-Burman languages, there is a mixture of the two orders of numeral and noun. Noun–numeral order is also the dominant type in New Guinea and among the languages of the easternmost islands of Indonesia.

Both types are interspersed amongst each other among the languages of the Pacific, including the eastern islands of New Guinea and in Australia. The two orders are about equally common among the languages of Canada and the United States east of the Rocky Mountains, with a notable pocket of noun–numeral languages in the southeastern United States.

Languages lacking a dominant order of numeral and noun are somewhat more common in Australia and in western Indonesia.

3 Theoretical issues

In many languages, numerals must occur with classifiers, as in (11).

(11) Nusu (Tibeto-Burman; southern China; Sun and Lu 1986: 44)

va^{53}*za*^{55}	*nga*^{55}	*gɟ*^{35}
little.pig	five	CLF

'five little pigs'

In such languages, the numeral is not really modifying the noun, at least not directly, but rather is modifying the classifier. For such languages, the map shows the order of the noun relative to this phrase consisting of numeral and classifier. Whether this phrase actually modifies the noun is itself something that is often unclear. It has been proposed for some languages that the classifier is the real head of the noun phrase and the numeral and noun are both dependents of the classifier. See Map 55 on numeral classifiers.

The issue of what element is the head also sometimes arises in languages without numeral classifiers. It has been proposed that some numerals in Russian serve as heads (Babby 1987). Similarly, in Rif Berber (Morocco), most modifiers follow the noun, while numerals precede; but the construction they occur in is the same as the genitive construction, as illustrated in (12), suggesting that the numeral is the head, not a modifier.

(12) Rif Berber (Kossmann 2000: 160, 108)

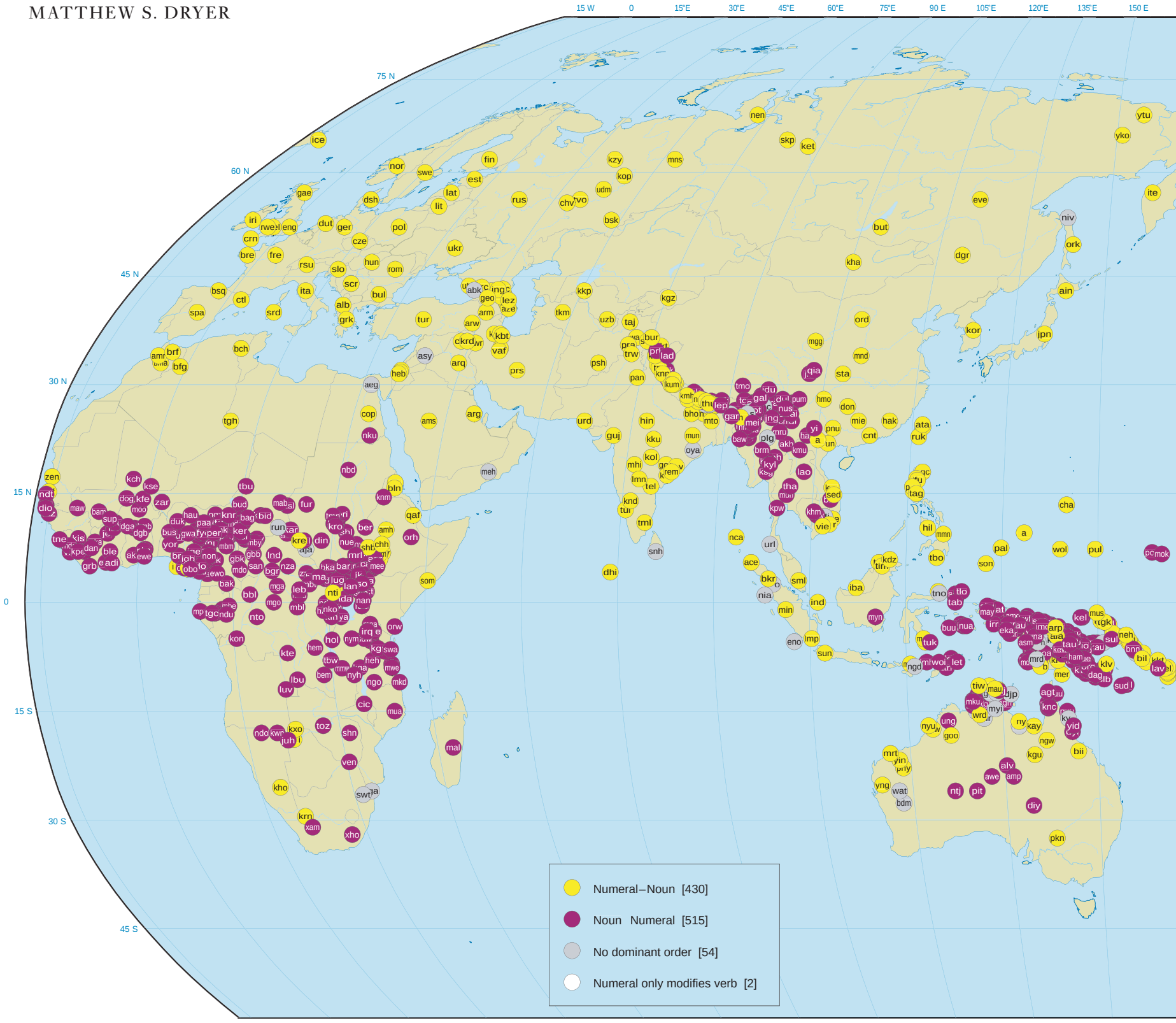
a. *t̲laṯa*	*n*	*təwrar*	b. *axxam*	*n*	*wəryaz*
three	GEN	hill	house	GEN	man
'three hills'			'the man's house'		

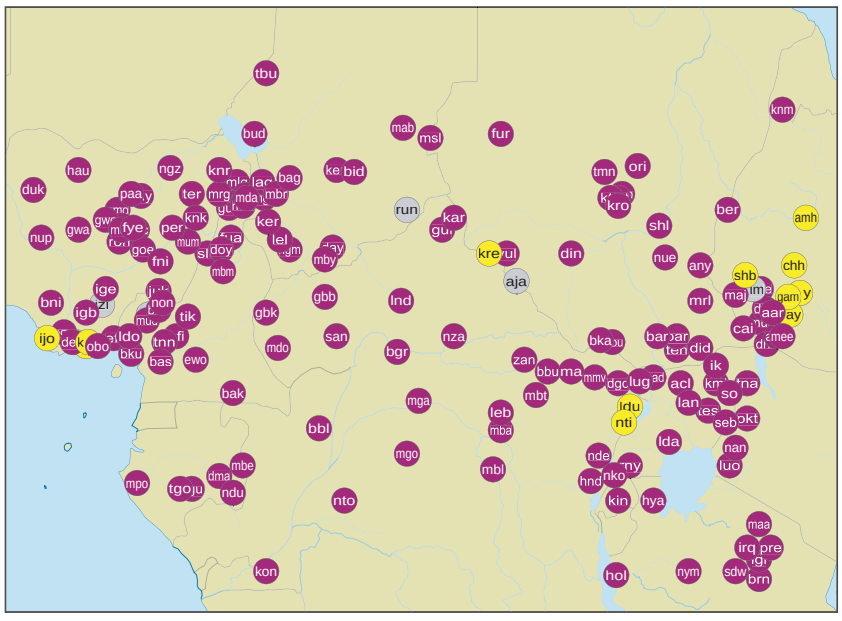
The issue of which element is the head is ignored in coding languages for this map.

There are other languages in which the numeral can be argued not to be a modifier of the noun. One type is represented by Kutenai (isolate; British Columbia, etc.), in which the numeral is the verb in an internally headed relative clause. This is illustrated above in (10a), where the noun *niu̯tahaɬ* 'young man' is grammatically the subject of the verb *ʔas* 'two'. For the purposes of this map, languages are coded according to the order of numeral and noun, regardless of whether the relationship is one of modification or not; since the normal order is that shown in (10a), Kutenai is coded as placing the numeral before the noun. The fact that the numeral is grammatically a verb in Kutenai (as it is also in Nias in (4) above) is ignored for the purposes of this map.

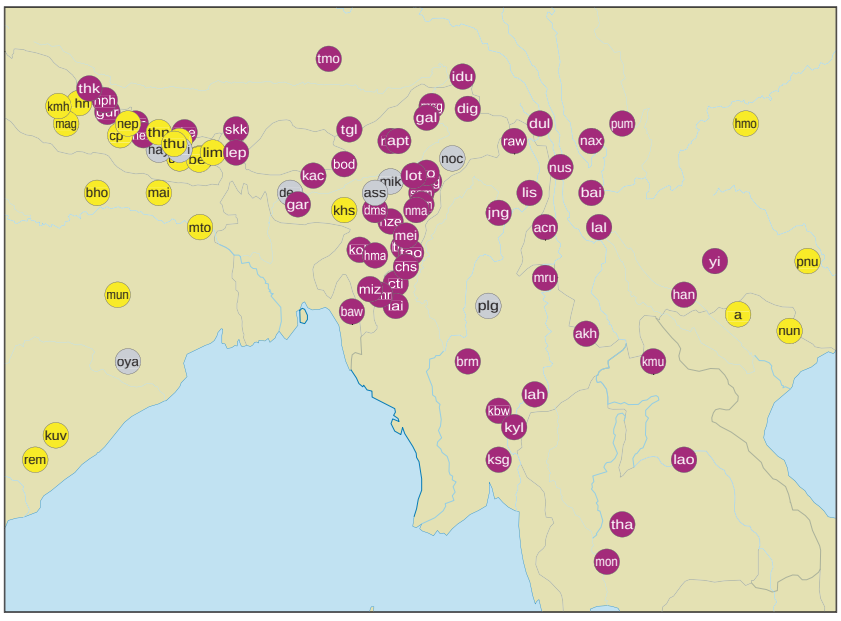
89 Order of Numeral and Noun

MATTHEW S. DRYER

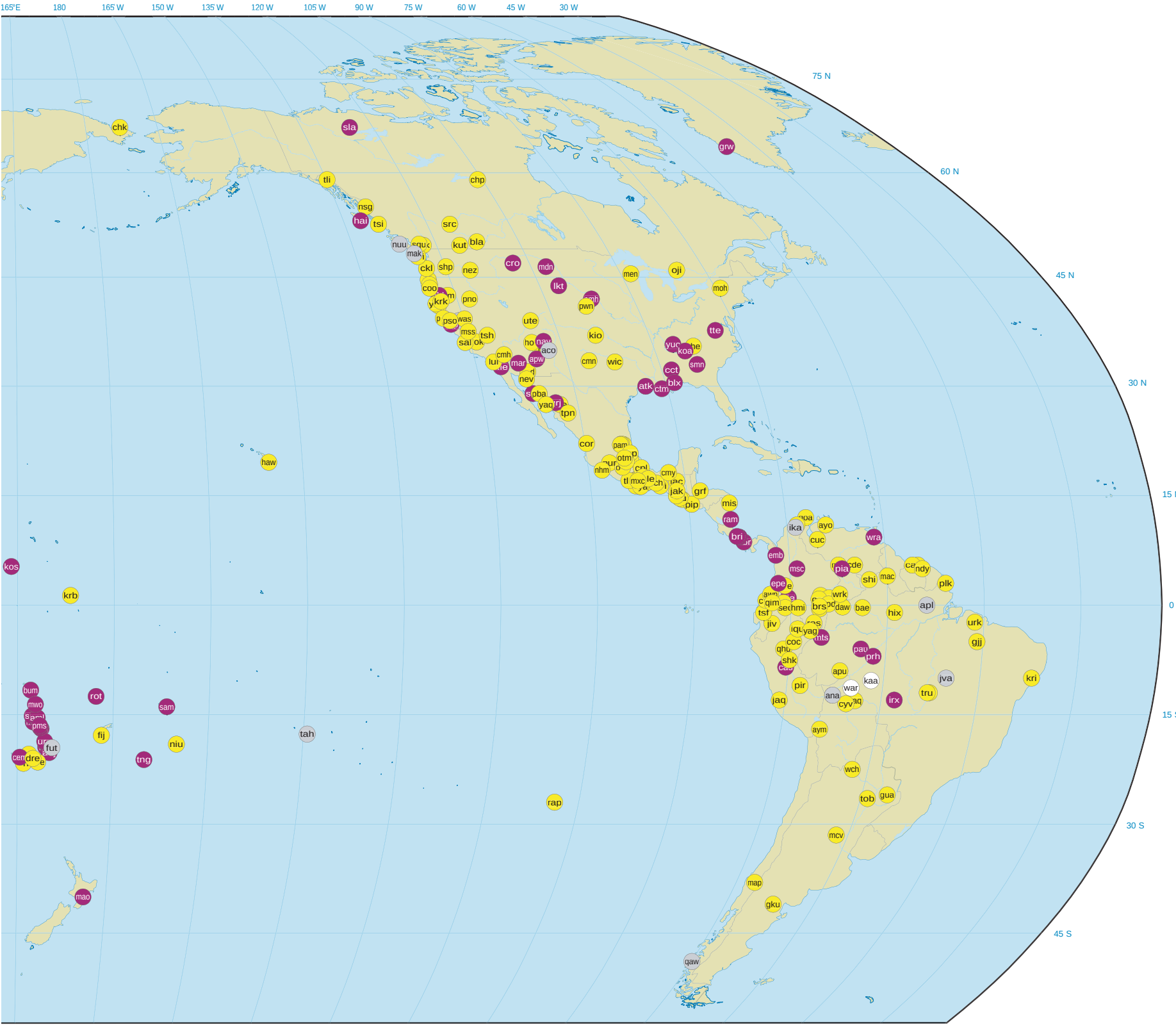


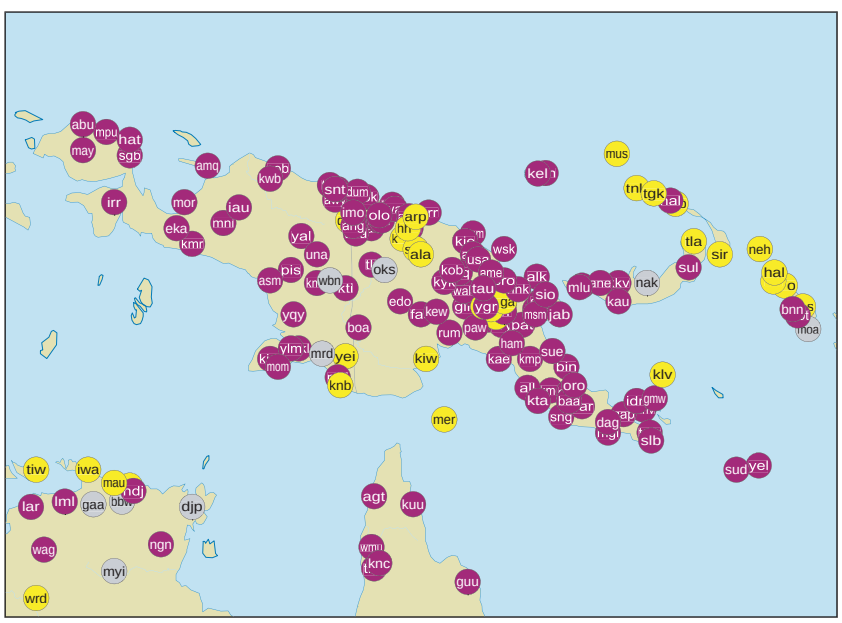


central Africa

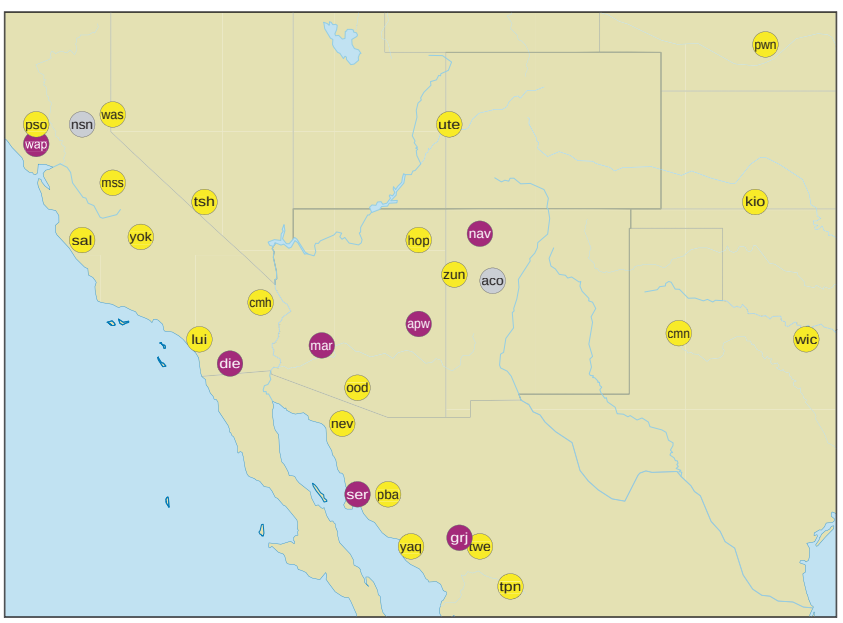


from Nepal to Thailand





New Guinea and northern Australia



southwestern United States and northern Mexico

90 Order of Relative Clause and Noun

MATTHEW S. DRYER

1 Defining the values

This map shows the order of relative clause and noun. A construction is considered a relative clause for the purposes of this map if it is a clause which, either alone or in combination with a noun, denotes something and if the thing denoted has a semantic role within the relative clause. If there is a noun inside or outside the relative clause that denotes the thing also denoted by the clause, that noun will be referred to as the **head** of the relative clause. Headless relative clauses (like English *what I bought at the store*) are not relevant to this map.

The two basic types shown on the map are languages in which **the relative clause follows the noun**, and languages in which **the relative clause precedes the noun**. The examples in (1) from English and from Maybrat (West Papuan) illustrate relative clauses following the noun.

(1) a. English
the book *[that I am reading]*
N Rel

b. Maybrat (Dol 1999: 137)
aof *[ro* *ana* *m-fat]*
sago [REL 3PL 3OBJ-fell]
'the sago tree that they felled'

The example in (2) from Alamblak (Sepik; Papua New Guinea) illustrates a relative clause preceding the noun.

(2) Alamblak (Bruce 1984: 109)
[ni *hik-r-fë]* *yima-r*
[2SG follow-IRREAL-IMMED.PST] person-3SG.M
Rel N
'a man who would have followed you'

1.	Relative clause follows noun (NRel)	507
2.	Relative clause precedes noun (RelN)	117
3.	Internally headed relative clause	18
4.	Correlative relative clause	7
5.	Adjoined relative clause	5
6.	Double-headed relative clause	1
7.	Mixed types of relative clause with none dominant	50
	total	705

The relative clauses illustrated in (1) and (2) occur with heads outside the relative clause; these can be referred to as **externally headed relative clauses**. In some languages, the head is inside the relative clause; these can be called **internally headed relative clauses**. These are illustrated by the examples in (3) from Mesa Grande Diegueño (Yuman; southern California and north-west Mexico); the fact that the head is inside the relative clause is clearest in (3a), in which the head (*gaat* 'cat') occurs between the subject and verb of the relative clause.

(3) Mesa Grande Diegueño (Couro and Langdon 1975: 187, 186)

a. *['ehatt* *gaat* *akewii]=ve=ch* *chepam*
[dog cat chase]=DEF=SUBJ get.away
'The cat that the dog chased got away.'

b. *['ehatt* *gaat* *kw-akewii]=ve=ch* *nye-chuukuw*
[dog cat REL.SUBJ-chase]=DEF=SUBJ 1OBJ-bite
'The dog that chased the cat bit me.'

What determines whether *'ehatt* 'dog' or *gaat* 'cat' is interpreted as the head in these examples is the presence versus absence of the subject relative prefix on the verb: its presence in (3b) signals that the head is the subject of the relative clause, namely *'ehatt* 'dog', while its absence in (3a) signals that the head is something other than the subject, in this case *gaat* 'cat'. Languages with internally headed relative clauses are probably more common than the map suggests because, until recently, grammarians often failed to recognize them as such.

The fourth type shown on the map is **correlative relative clauses**, as in (4) from Bambara (Mande, Niger-Congo; Mali).

(4) Bambara (Bird and Kante 1976: 9)
[muso *min* *taara],* *o* *ye* *fini* *san*
[woman REL leave] 3SG PST cloth buy
'The woman who left bought the cloth.'

Correlative clauses are strictly speaking a subtype of internally headed relative clauses in that the head noun occurs inside the clause, but they differ from those coded here as internally headed in that the relative clause is outside the main clause and is connected anaphorically to a noun phrase in the main clause that corresponds to the head noun in the English translations.

The fifth type shown on the map consists of languages with **adjoined relative clauses**. As with the preceding type, adjoined relative clauses are outside the main clause; they do not form a constituent with the head noun, which is in the main clause, and they may be separated from it. However, unlike correlative clauses, the head occurs in the main clause rather than in the relative clause. An example of a language in which relative clauses are of this sort is Diyari (Pama-Nyungan; South Australia), as illustrated in (5) (Austin 1981: 188).

(5) Diyari (Austin 1981: 210)
ŋaṉi *wiḻa-ṉi* *yaṭa-ḻa* *ŋana-yi*
1SG.SUBJ woman-LOC speak-FUT AUX-PRES
[yinda-ṉaṉi]
[cry-REL.DS]
'I'll talk to the woman who is crying.'

The sixth type shown on the map is represented by a single language, Kombai (Trans-New Guinea; Papua, Indonesia), and is referred to here as a **double-headed relative clause**. As illustrated in (6), relative clauses in Kombai combine the features of externally headed and internally headed relative clauses in a single structure: they have both an external head noun and a noun corresponding to the head noun inside the relative clause. While the two nouns are sometimes the same, as in (6a), the external noun is usually more general than the one inside the relative clause, as in (6b), where the external noun is simply *ro* 'thing'.

(6) Kombai (de Vries 1993: 78, 77)

a. *[doü* *adiyano-no]* *doü* *deyalukhe*
[sago give.3PL.NONFUT-CONN] sago finished.ADJ
'The sago that they gave is finished.'

b. *[gana* *gu* *fali-kha]* *ro*
[bush.knife 2SG carry-go.2SG NONFUT] thing
'the bush knife that you took away'

The final type shown on the map involves languages which employ two or more of the preceding constructions, without one being dominant. For example, Gimira (Omotic; Ethiopia) allows the relative clause to either precede or follow the head noun, and there is no evidence for one order being dominant (Breeze 1990: 39). In most languages that allow both orders of relative clause and noun, there appear to be reasons for treating one as dominant (see the box section "Determining Dominant Word Order" on p. 371). In Kapampangan (Austronesian; Philippines), for example, relative clauses can precede the noun only if they consist of a single word,

whereas relative clauses of any length can follow the noun (Mirikitani 1972: 189). Kapampangan is thus shown on the map as having postnominal relative clauses. Some languages have both externally headed and internally headed relative clauses or correlative clauses, without one type being dominant. For example, Murrinh-Patha (isolate; Northern Territory, Australia) has both postnominal relative clauses, as in (7a), and internally headed relative clauses, as in (7b).

(7) Murrinh-Patha (Walsh 1976: 289, 287)

a. *ŋayi-e* *nʸinʸi* *ŋu-nʸi-baḍ-nu* *tʸu*
1SG-ERG 2SG.ABS 1SG-2SG-strike-FUT weapon
kumukur-te *[yile* *ŋayi* *mam-ŋa-wata*
club-INSTR [father 1SG 3SG.PERF-1SG.BEN-make
ŋayi-nu]
1SG-DAT]
'I will hit you with the club my father made for me.'

b. *[mutyiŋga-e* *ŋayi* *pan-ŋi-baḍ]*
[woman-ERG 1SG.ABS 3SG.PERF-1SG-hit]
paŋanduwi *mundakŋayya*
3SG.PERF.arrive earlier
'The old woman who hit me arrived earlier.'

That the head *kumukur* 'club' in (7a) is external is clear from the fact that it bears an instrumental suffix indicating its role in the main clause, not the relative clause (where it would be in absolutive case). That the head *mutyiŋga* 'woman' in (7b) is internal to the relative clause is clear from the fact that it bears ergative case, reflecting its role in the relative clause, not its role in the main clause. Other examples of languages having more than one relative construction with none dominant include Panjabi, which has both prenominal relative clauses and a correlative construction (Bhatia 1993: 50, 53, 56), Ngiyambaa (Pama-Nyungan; New South Wales, Australia), which has both a correlative construction and an adjoined relative construction (Donaldson 1980: 297–9), and Kobon (Trans-New Guinea; Papua New Guinea), which has both prenominal and double-headed relative clauses (Davies 1981: 29).

Note that languages vary as to whether relative clauses involve forms of verbs that also occur in main clauses. In English relative clauses like that in (1a), for example, the forms of the verbs are the same, as illustrated by *am reading* in (1a). In Kolyma Yukaghir (isolate; north-east Siberia), in contrast, the verb in relative clauses lacks the inflections associated with finite verbs and occurs with a general attributive suffix, as in (8).

(8) Kolyma Yukaghir (Maslova 2003: 418)
[purk-in *šoromo* *le:-je]* *šoromo*
[seven-ATTR person eat-ATTR person
'a person who has eaten seven people'

English also has nonfinite participial relative clauses, as in *the man reading the book*. Nonfinite relative clauses are sometimes not considered as relative clauses; however, since there are many languages where relative clauses are all nonfinite and since these constructions mean the same thing as finite relative clauses in English, such participial constructions are considered as relative clauses here.

In some languages, headless relative clauses are arguably the basic form of relative clauses, and relative clauses with a head can be analysed as nominal expressions in apposition to the head. Such an analysis is argued for by Curnow (1997) for Awa Pit (Barbacoan; Colombia and Ecuador); (9a) illustrates a headless relative clause, while the example in (9b) illustrates a relative clause followed by a head noun.

(9) Awa Pit (Curnow 1997: 286)

a. *na=na* *[pishkatu* *pay-nin-tu=mika]=ta*
1SG=TOP [fish buy-CAUS-IMPF=NMLZ.SG]=ACC
pyan-ta-w
hit-PST-1
'I hit the one who was selling the fish.'

b. *[na-wa* *sula* *kwa-t=mika]*
[1SG-ACC bite bite-PERF.PTCP=NMLZ.SG]
kwizha *nya* *walkwa-ti-zi*
dog meat steal-PST-2/3
'The dog which bit me stole the meat.'

The fact that the headless relative clause in (9a) is itself a nominal expression is further brought out by the fact that it occurs with the postpositional accusative clitic *=ta*. Relative clauses like these are often called nominalizations, but are still considered relative clauses for the purposes of this map.

2 Geographical distribution

The geographical distribution of the two major types, i.e. externally headed relative clauses with relative clauses preceding or following the noun, is quite clear. The overwhelmingly dominant type in much of the world is for the relative clause to follow the noun. The exception to this is much of Asia, where except in the Middle East and South-East Asia, the dominant type has the relative clause preceding the noun. Outside Asia, except for a scattering of geographically isolated instances, this prenominal type is found in only three relatively small areas: (i) New Guinea; (ii) Ethiopia and Eritrea; and (iii) north-western South America. However, even in these three areas, languages with prenominal relative clauses are a minority.

Internally headed relative clauses are scattered throughout the world, though two areas where they are more common are worth mentioning. One is North America: though they constitute only a minority here, they are more common than externally headed relative clauses preceding the noun. The other is a small area in West Africa where both internally headed and correlative clauses are common. A second area where correlative clauses are found is South Asia. It must be stressed that the map only shows languages with a specific type if that type is dominant in the language. There are more languages with internally headed or correlative clauses than the map might suggest, because these often co-exist with some other type and are not dominant in the language.

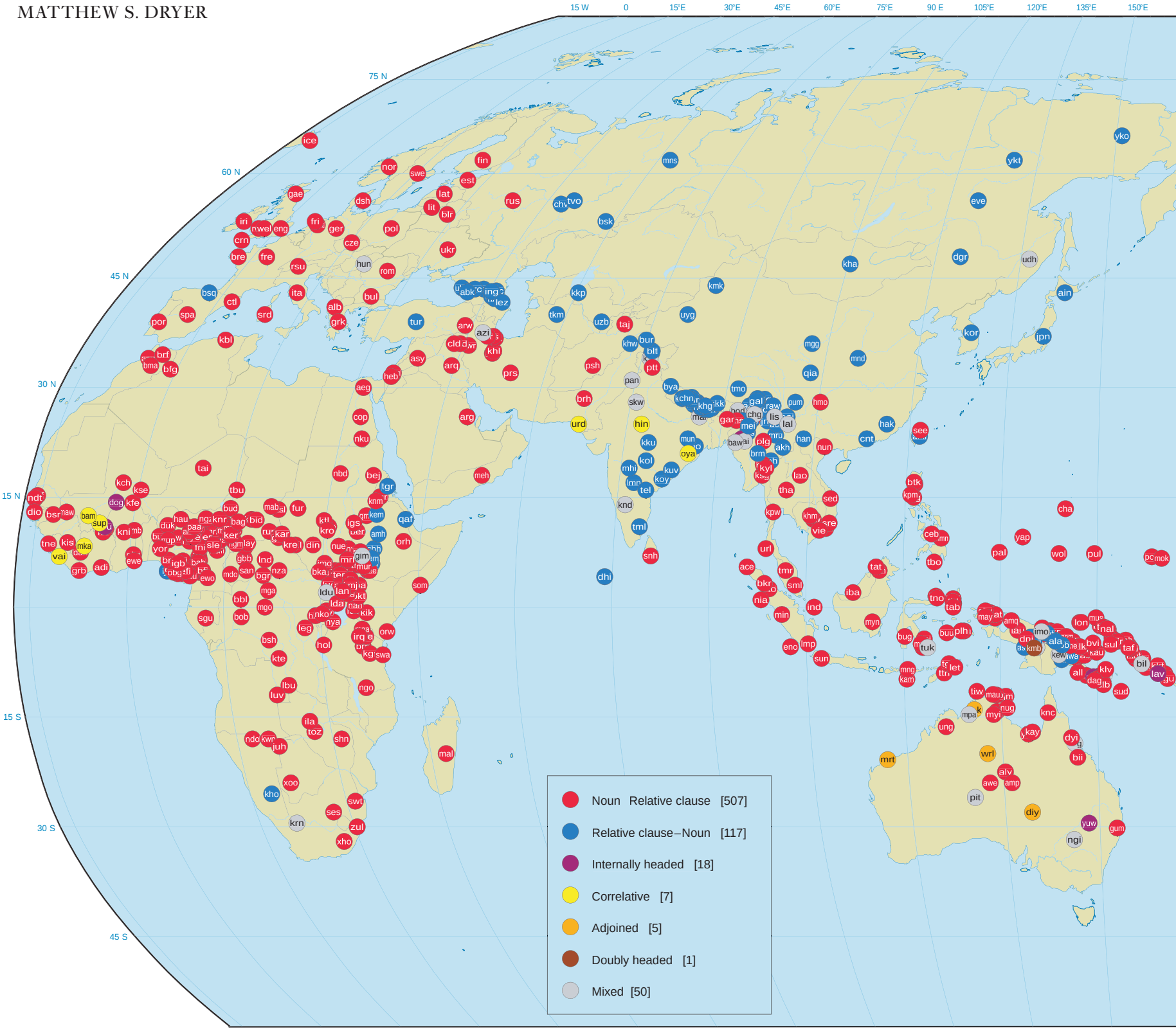
There are only five languages shown on the map as employing adjoined relative clauses as the dominant type, four of them in Australia and one in South America. Only one language is shown with double-headed relative clauses as the dominant type, in Papua (Indonesia).

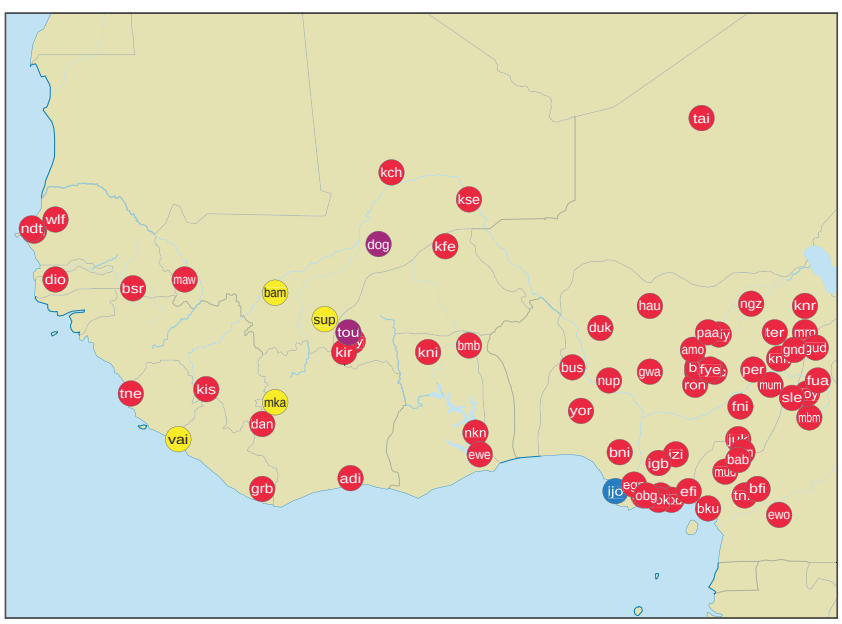
3 Theoretical issues

The concentration of languages in which the relative clause precedes the noun in much of Asia, but with low frequency elsewhere in the world, illustrates the extent to which a region as large as Asia can be a linguistic area. In fact, this area is more generally associated with a type in which all modifiers precede the noun, a type that is relatively uncommon outside of Asia. While most of the languages with this property are verb-final languages, it is clear that this type is found only among a minority of verb-final languages elsewhere in the world. The relationship between the order of relative clause and noun and the order of object and verb is discussed further in Chapter 96. Maps 122 and 123 show other features involving relative clauses.

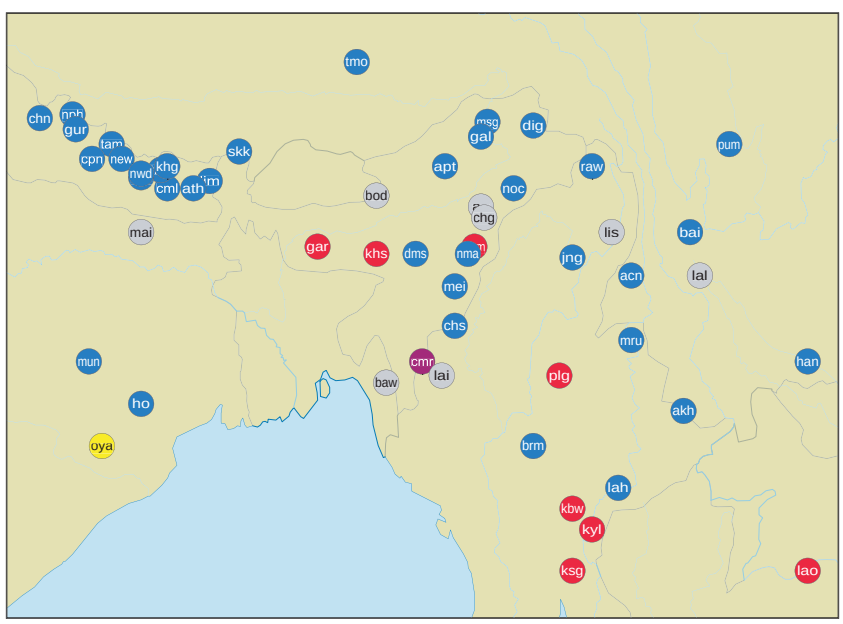
90 Order of Relative Clause and Noun

MATTHEW S. DRYER

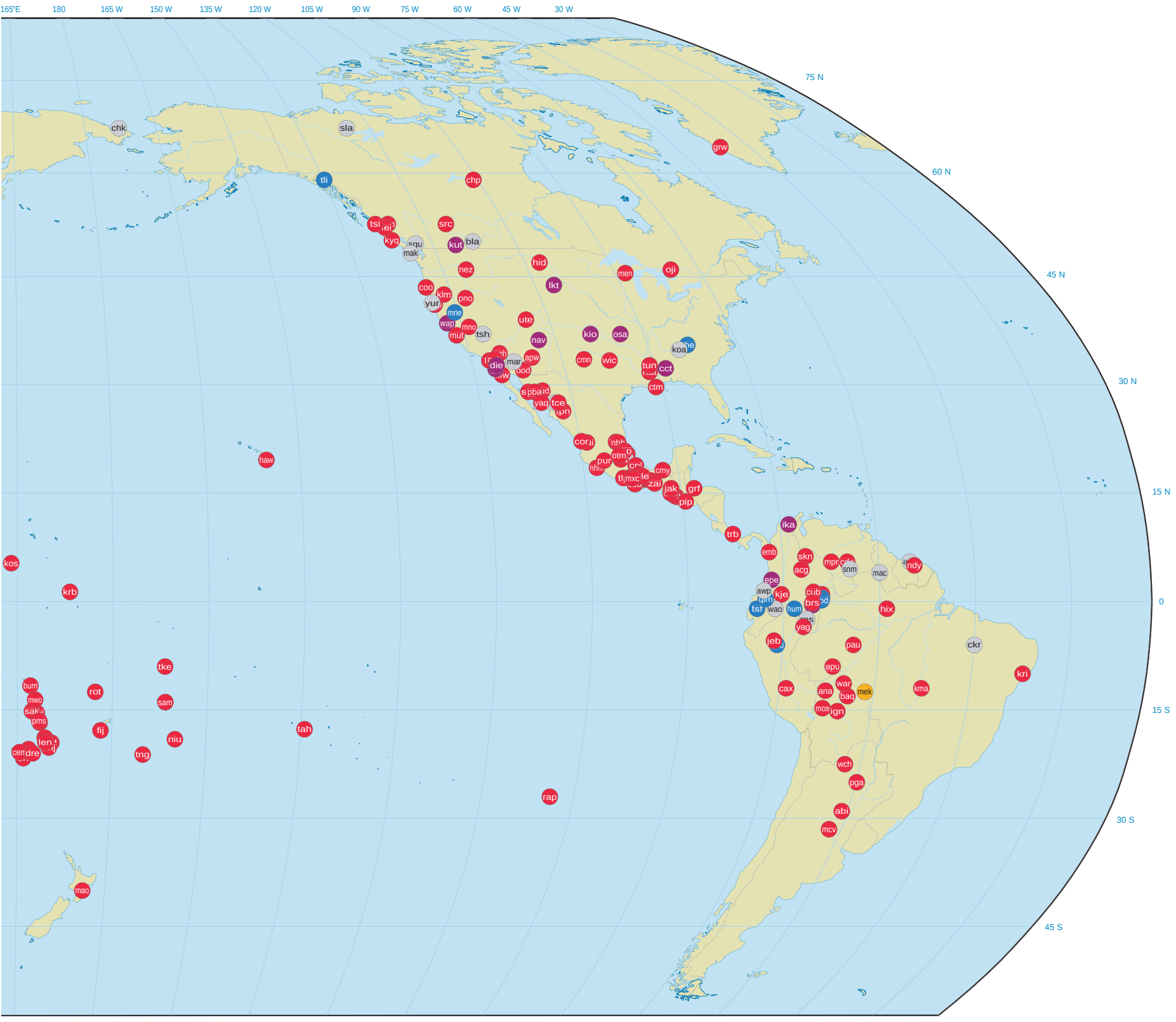


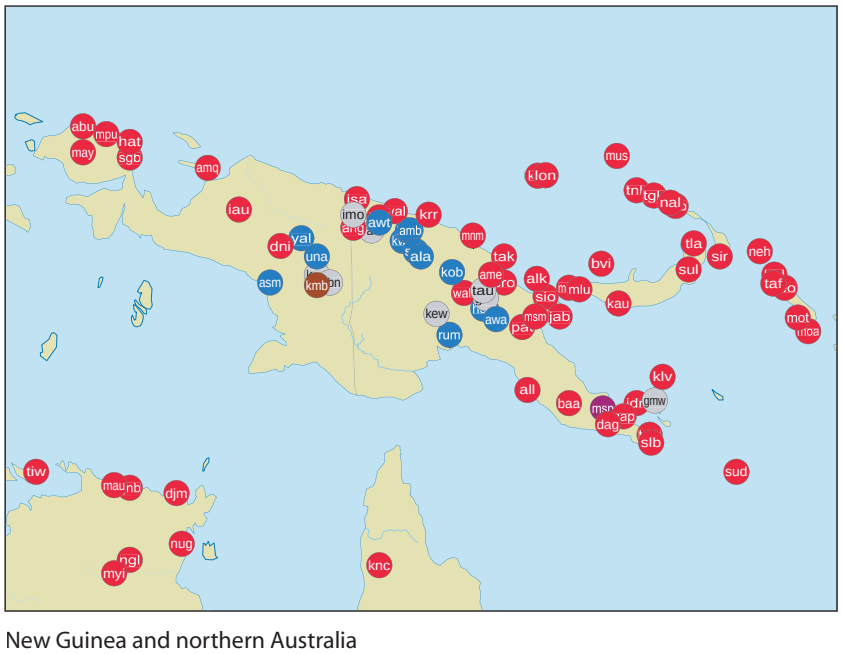


West Africa

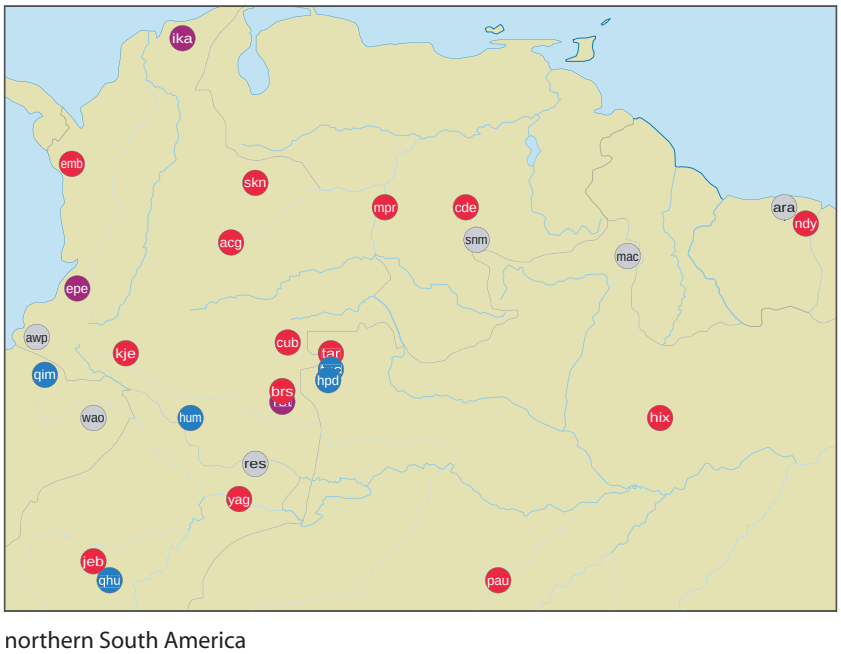


Nepal, eastern India, and Myanmar





New Guinea and northern Australia



northern South America

91 Order of Degree Word and Adjective

MATTHEW S. DRYER

1 Defining the values

This map shows the position of degree words with respect to the adjective that they modify. For the purposes of this map, the term *adjective* should be interpreted in a purely semantic sense, as a word denoting a property, since in many languages the words in question do not form a separate word class, but are verbs or nouns. Degree words are words with meanings like 'very', 'more', or 'a little' that modify the adjective to indicate the degree to which the property denoted by the adjective obtains. Degree words are traditionally referred to as adverbs, though in many languages the degree words do not belong to the same word class as adverbs; even for English there is little basis for saying that degree words belong to the same word class as adverbs which modify verbs.

1. Degree word precedes adjective (DegAdj)	205
2. Degree word follows adjective (AdjDeg)	177
3. Both orders occur with neither order dominant	55
total	437

The first type shown on the map consists of languages in which **the degree word precedes the adjective**. An example of such a language is Pumi (Tibeto-Burman; China), as in (1).

(1) Pumi (Ding 1998: 107)
lealián *ggáo*
very deep
Deg Adj
'very deep'

Most European languages, like English (e.g. *very tall*, *too small*, *somewhat afraid*), are also instances of this type. The second type consists of languages in which **the degree word follows the adjective**, as in the example in (2) from Kairiru (Oceanic; Papua New Guinea).

(2) Kairiru (Wivell 1981: 74)
nau *pulau* *sek*
sea murky too
Adj Deg
'. . . the sea is too murky.'

The third value shown on the map includes languages in which **both orders occur with neither dominant** (see the box section "Determining Dominant Word Order" on the next page). In some languages, such as Kisi (Atlantic, Niger-Congo; Guinea), degree words in general can either precede or follow the adjective (Childs 1995: 256). In many other languages, however, individual degree words differ as to whether they precede or follow the adjective. For example, in Wari' (Chapacura-Wanhan; Brazil), the degree word meaning 'a little' precedes the adjective, as in (3a), while the word meaning 'very' follows, as in (3b); no other degree words are apparently mentioned by Everett and Kern (1997).

(3) Wari' (Everett and Kern 1997: 346)
a. *'amon* *mixem*
a.little black
'a little dirty'
b. *mixem* *tamana*
black very
'very dirty'

In English, the degree word *enough* differs from other degree words in that it follows the adjective (*large enough* versus *very large*). One order is considered dominant if the number of degree words that occur on one side of the adjective is more than twice the number that occur on the other side. For example, in Indonesian, Sneddon (1996: 177–81) lists fifteen degree words that precede the noun, four that follow, and one that either precedes or follows, so Indonesian is coded on the map as placing the degree word before the adjective. Conversely, Wari' is coded as a language with both orders where neither order is dominant.

In some languages, the order of degree word and adjective depends on whether the adjective is being used attributively, i.e. modifying a noun, or predicatively. This is the case in Ndyuka (Creole; Suriname), in which a degree word precedes an adjective used attributively, as in (4a), but follows an adjective used predicatively, as in (4b).

(4) Ndyuka (Huttar and Huttar 1994: 173, 175)
a. *wan* *tumisi* *gaan* *makiti*
INDEF very great power
'a very great power'
b. *i* *kon* *dyendee* *tumisi*
2SG come elegant very
'You've become very elegant.'

Tetelcingo Nahuatl (Tuggy 1979: 76) and Quiotepec Chinantec (Robbins 1968: 59–60) are similar.

In some languages, there are degree morphemes which occur as affixes on adjectives. The comparative and superlative suffixes in English (*-er* in *stronger*, *-est* in *strongest*) are examples of degree affixes. Similarly, in many languages, the meaning 'very' is expressed by an affix, as in the example in (5) from Maricopa (Yuman; Arizona).

(5) Maricopa (Gordon 1986: 141)
man-sh *m-hmii-hot-m*
2SG-SUBJ 2.SUBJ-tall-very-REALIS
'You are very tall.'

This map does not include degree affixes, restricting attention to separate words expressing degree. Some languages are not included on the map because the only degree morphemes mentioned in available descriptions are affixes.

2 Geographical distribution

Languages in which the degree word precedes the adjective constitute the overwhelmingly dominant type, with very few exceptions, in Europe and Asia, except in the Middle East and South-East Asia. It is also the dominant type in North America, though with more exceptions. In South America it is the dominant type along the western side of the continent. It is a minority type in Africa and New Guinea. The distribution of types is quite mixed in Australia and among the Austronesian languages of Indonesia, the Philippines and the Pacific.

Languages in which the degree word follows the adjective are the dominant type in Africa and in New Guinea. They are the dominant type in South America except down the western side of the continent. They also constitute the dominant type in the mainland of South-East Asia, in an area extending westward to include languages along the border between India and Myanmar.

Languages lacking a dominant order are widespread but are distinctly infrequent in much of Europe and Asia, again with the exception of the Middle East and South-East Asia. They are particularly common in South-East Asia and among Austronesian languages. The strong areal pattern across much of Europe and Asia is in striking contrast to the lack of patterning within Austronesian languages. Even within relatively small regions, such as Sumatra or the Philippines, both orders are found.

DETERMINING DOMINANT WORD ORDER

Matthew S. Dryer

A number of maps in this atlas show the **dominant word order** of various sets of elements, in most cases pairs of elements (like adjective and noun) but in some cases sets of three elements (like subject, object, and verb). For any set of elements, there are some languages in which only one order is permitted and other languages in which more than one order is permitted. Among languages of the latter sort, one can further distinguish languages in which one order is used more frequently than others from languages in which this is not the case. For example, with respect to the order of adjective and noun, there are languages which only employ adjective–noun order, others that only employ noun–adjective order, and still others that allow both orders. Among languages that allow both orders, there are some in which adjective–noun order is more frequent, some in which noun–adjective is more frequent, and some in which both orders occur with comparable frequency.

Where a language is shown on one of the word-order maps as having a particular order as the dominant order in the language, this means that it is either the **only order possible** or the order that is **more frequently used**. The maps do not distinguish these two possibilities, because it is often not possible to obtain reliable information from descriptive grammars on whether a particular order which is not the most frequent order is grammatical or not. While a grammar may say, for example, that the order of adjective and noun in a language is adjective–noun, it often turns out that the alternate order is possible, either in special discourse contexts or in special grammatical contexts, so it is rarely possible to conclude with confidence that only one order is permitted.

The expression "dominant order" is used here, rather than the more common expression "basic order", to emphasize that priority is given here to the criterion of what is more frequent in language use, as reflected in texts. The reason for assigning priority to this criterion is that for most languages, this is the only criterion for which we have any relevant information. When a language allows both orders of adjective and noun, for example, grammars will often mention this but describe one order as the normal order or the more frequent order. For some languages, the classification of a language in this atlas is based on actual text counts. The rule of thumb employed is that if text counts reveal one order of a pair of elements to be more than twice as common as the other order, then that order is considered dominant, while if the frequency of the two orders is such that the more frequent order is less than twice as common as the other, the language is treated as lacking a dominant order for that pair of elements. For sets of three elements, one order is considered dominant if text counts reveal it to be more than twice as common as the next most frequent order; if no order has this property, then the language is treated as lacking a dominant order for that set of elements. Of course, unless one examines a large number and a broad variety of texts, one cannot be sure that differences in frequency may not occasionally reflect the idiosyncratic properties of a particular set of texts. It is likely that in some cases, further text counts would lead to classifying a language differently.

For some languages, the classification on the map is based on a claim in the source that some order is basic or that it is pragmatically neutral. In the absence of evidence to the contrary, I assume that these are also the dominant orders. Occasionally, however, such claims are at odds with frequency data provided by the author. For example, Abbott (1991: 25) characterizes OVS order (object–verb–subject) in Macushi (Carib; Brazil) as basic, and says that SOV order is used to highlight the subject. However, she cites text count data that show that OVS and SOV order are about equally common. I base my classification of Macushi here on the frequency counts, and since no order is more than twice as frequent as the next most frequent order, I treat this language as lacking a dominant order of subject, object, and verb.

For some word-order features where more than one order is possible, such as the order of object and verb, the order will generally be determined syntactically or by extragrammatical factors. But for other word-order features, it may be largely determined by specific lexical items. For example, in languages with both prepositions and postpositions, it is generally the case that each adposition is either always a preposition or always a postposition. In such cases, the classification of a language as prepositional or postpositional is based here on a combination of whether the number of prepositions outnumbers the number of postpositions (or vice versa) and which adpositions express basic meanings and are thus likely to be used more frequently. For example, in Koyraboro Senni (Songhay; Mali), there are over a dozen postpositions but only three prepositions, and among the postpositions are a number with apparently higher frequency of usage, including one marking indirect objects, one marking locatives (covering meanings of 'at', 'to', or 'from'), and one meaning 'on', while the prepositions tend to have more specialized meanings ('since', 'until', and 'during', though also 'with') (Heath 1999a). Because this suggests that postpositions are more frequent, Koyraboro Senni is classified here as postpositional.

Similarly, in Korowai (Trans-New Guinea; Papua, Indonesia), all adjectives can precede the noun, as in (1a), but a few, like the adjective meaning 'big', can also follow the noun, as in (1b).

(1) Korowai (van Enk and de Vries 1997: 69)

a. *lembul* *nggulun*
bad teacher
'a bad teacher'

b. *yanop* *khonggél-khayan*
man big-very
'a very big person'

Again, it is assumed from this description that adjective–noun order is more frequent, and Korowai is treated on Map 87 as adjective–noun.

Some grammars will describe a particular word order as more contrastive. It is assumed from statements of this sort that the more contrastive order is used less frequently; hence the language will be coded according to the noncontrastive order. For example, in Asmat (Trans-New Guinea; Papua, Indonesia) adjective–noun order is described as contrastive, while noun–adjective order is neutral, as in (2).

(2) Asmat (Voorhoeve 1965b: 140)

a. *ów* *akát*
people good
'good people'

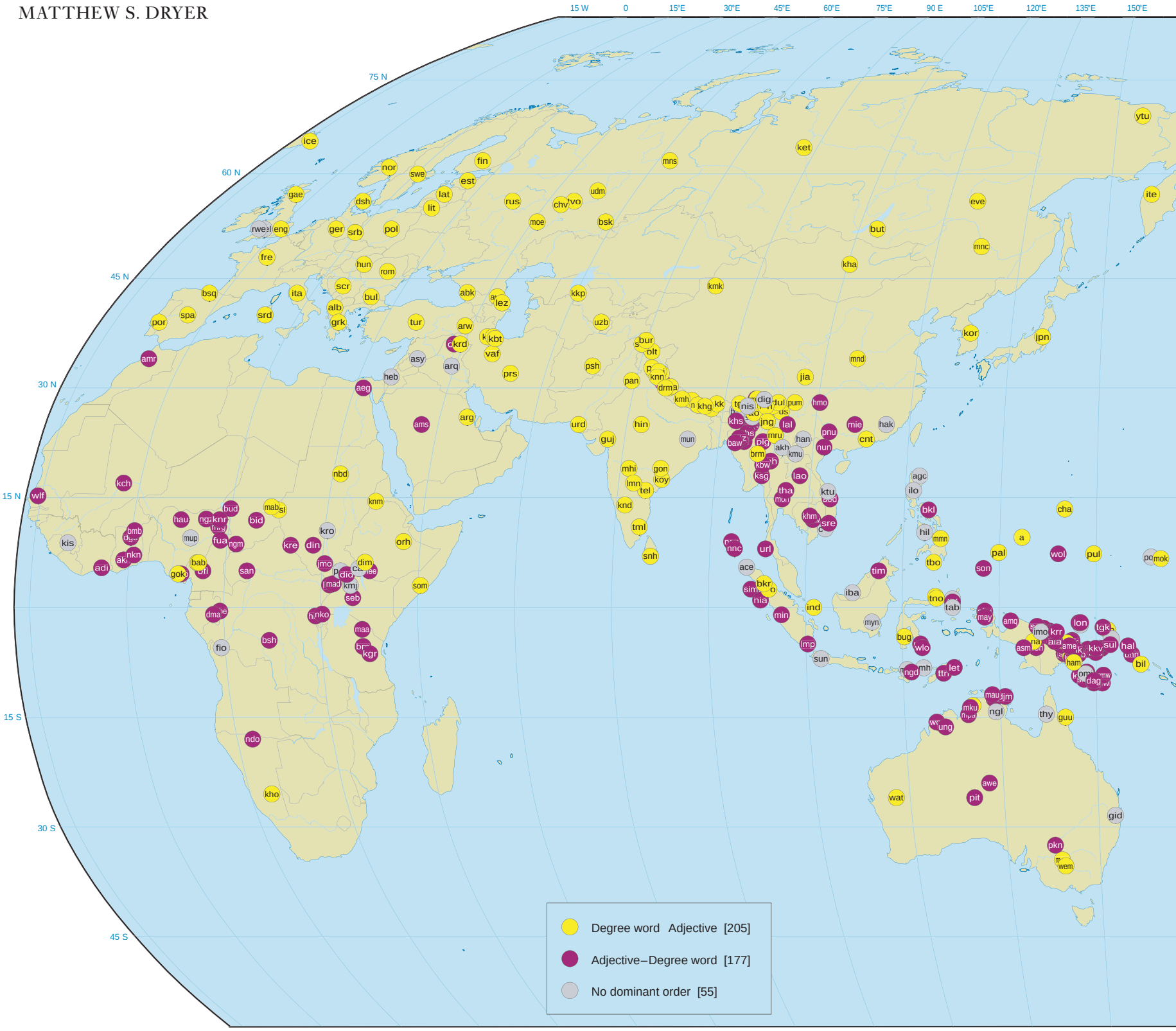
b. *akát* *ów*
good people
'good people (in contrast to bad people)'

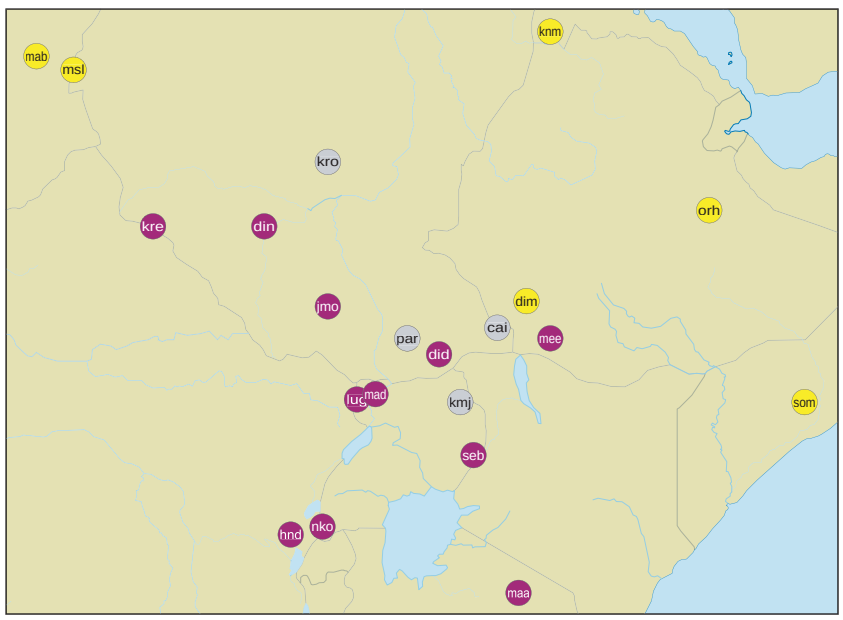
The situation is similar in Ilocano (Austronesian; Philippines) except that the situation is reversed (Rubino 1998): in Ilocano, the neutral order is adjective–noun, while noun–adjective order is contrastive.

If a grammar indicates that both orders of a pair of elements are possible, without stating that one is more common or without any comment suggesting that one order is more common, then the language will be shown on the map as having both orders without one being dominant.

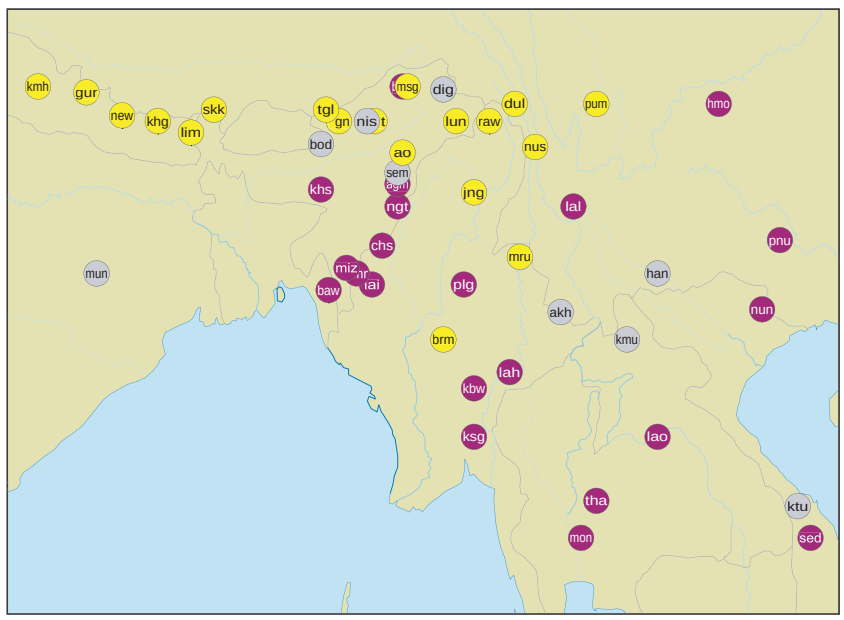
91 Order of Degree Word and Adjective

MATTHEW S. DRYER

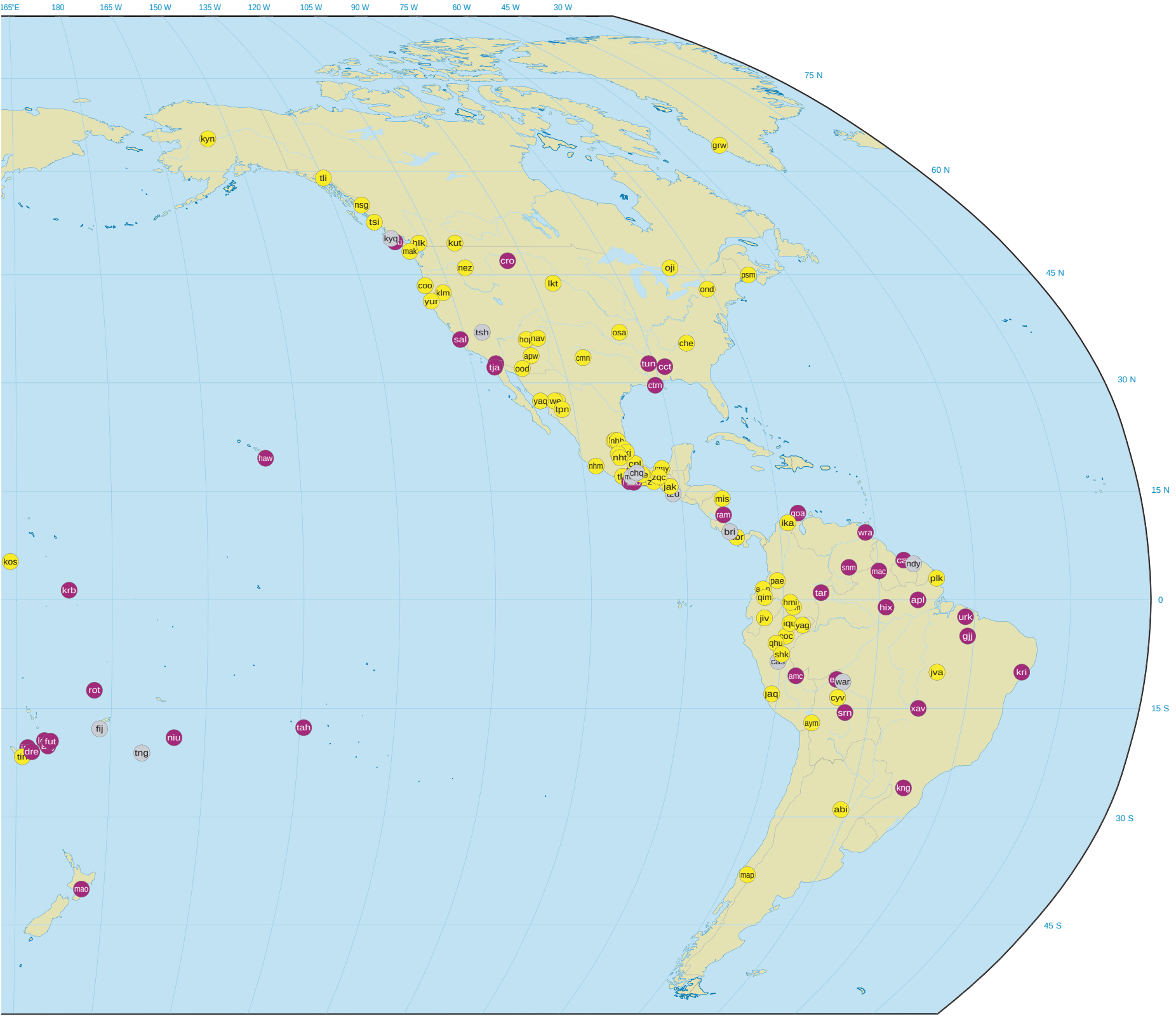


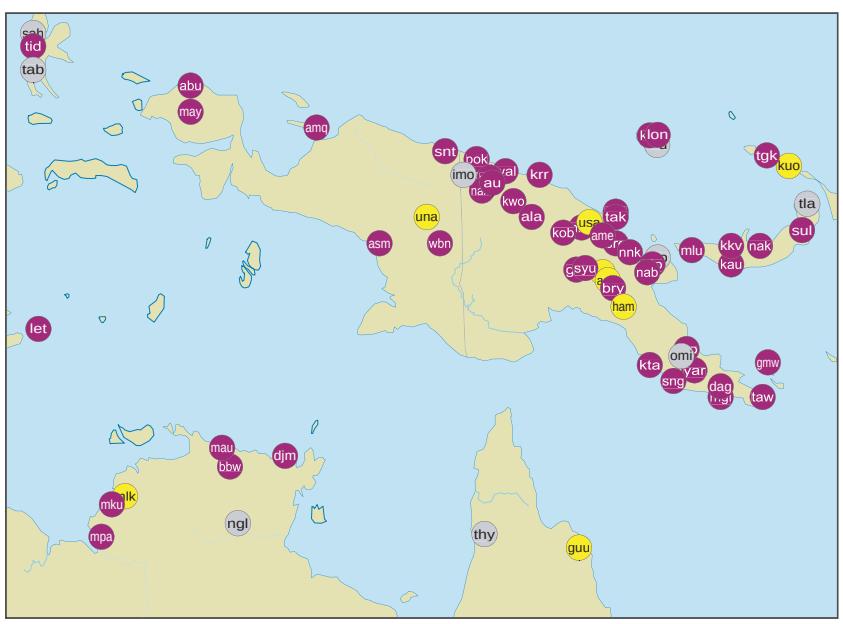


eastern Africa

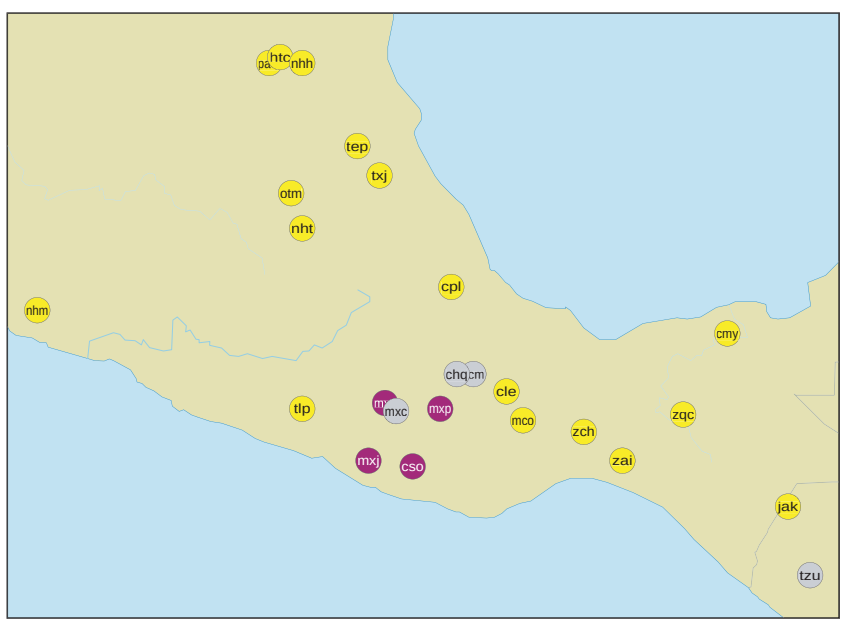


Myanmar and surrounding areas





New Guinea and northern Australia



Mesoamerica

92 Position of Polar Question Particles

MATTHEW S. DRYER

1 Defining the values

This map shows the position of question particles in polar questions. Polar questions are ones that elicit the equivalent of a 'yes' or 'no' response. This map does not include interrogative affixes on verbs, although it should be noted that the distinction between interrogative affixes and separate interrogative particles is often hard to make. The map treats interrogative clitics that can attach to words of different categories as question particles. Map 116 shows the distribution of different ways to signal polar questions, including question particles, but does not show their position in the sentence. A number of languages shown on Map 116 as having question particles are not shown on this map because their position is not clear from available materials.

1. Question particle at beginning of sentence	118
2. Question particle at end of sentence	272
3. Question particle in second position in sentence	45
4. Question particle with other position	8
5. Question particle in either of two positions	24
6. No question particle	310
total	777

In some languages, the question particle occurs **at the beginning of the sentence**, as in (1) from !Xóõ (Southern Khoisan; Botswana).

(1) !Xóõ (Traill 1994: 18)
lú tûu à sîi
Q people TNS come
'Did the people come?'

The French expression *est-ce que*, illustrated in (2), though originally a verb plus demonstrative plus complementizer, can be considered to have become grammaticized as a question particle; French is thus treated here as a language with an initial question particle.

(2) French (Harris 1988: 237)
Est-ce que *le président vient?*
Q the president come.PRES.3SG
'Is the president coming?'

In other languages, the question particle occurs **at the end of the sentence**, as in (3) from Hatam (West Papuan; Indonesia).

(3) Hatam (Reesink 1999: 69)
a-yai bi-dani mem di-ngat ***i***
2SG-get to-me for 1SG-see Q
'Would you give it to me so that I can see it?'

Sentence-final question particles often cliticize onto the last word in the sentence, as in (4) from Majang (Surmic, Nilo-Saharan; Ethiopia).

(4) Majang (Unseth 1989: 126)
*dɛn-ɛ daaki tolay=**ŋ***
see-3SG Daaki Tolay=Q
'Did Daaki see Tolay?'

Occasionally a question particle will code other grammatical features of the clause. For example, in Aghu (Trans-New Guinea; Papua, Indonesia), the question particles also code tense (Boelaars 1950: 152).

A third common position for question particles is **second position in the sentence**, following the first word or constituent, as in (5) from Yurok (Algic; California).

(5) Yurok (Robins 1958: 139)
kic ***hes*** *nesk^{W}ec-ok'^{W} ku wʌʔyʌs*
PST Q come-3SG DEF girl
'Has the girl come back yet?'

Second-position question particles very often encliticize onto the first constituent, as in the example in (6) from Mono (Uto-Aztecan; California).

(6) Mono (Norris 1986: 44)
*Charley=**m̃a?*** *mia-pɨ*
Charley=Q go-PERF
'Has Charley left?'

The fourth type shown on the map consists of languages in which there is a question particle with a specific position within the clause that is something **other than initial position, second position, or final position**. For example, in Hmong Njua (Hmong-Mien; China), the question particle occurs after the subject and before the verb, as in (7).

(7) Hmong Njua (Harriehausen 1990: 205)
kuv cov nyaaj ***puas*** *nyam hoob pee*
1SG MEAS aunt Q like sing song
'Do my aunts like to sing?'

In Niuean (Oceanic; Niue Island, Pacific), which is VSO, the question particle immediately follows the verb, as in (8).

(8) Niuean (Seiter 1980: 25)
foaki age ***nakai*** *e koe e fakamailoga*
give DIRL Q ERG you ABS prize
ki a taha
to PERS someone
'Did you give the prize to someone?'

The fifth type shown on the map is languages in which question particles can occur in **two of the positions** represented by the first four types, without one position being dominant. This includes languages where question particles occur either at the beginning or at the end of the sentence. An example of such a language is Koyraboro Senni (Songhay; Mali), in which the question particle is the same as the word meaning 'or' (Heath 1999a: 225). This type also includes languages with two different question particles that occur in different positions. An example of such a language is Hunde (Bantu; Democratic Republic of Congo), in which one question particle occurs sentence-initially, as in (9a), and the other sentence-finally, as in (9b).

(9) Hunde (Kahombo 1992: 171)
a. ***mbéni*** *ámukátsí mu-lómbe*
Q woman NC-lazy
'Is the woman lazy?'
b. *ámukátsí mu-lómbe* ***hé***
woman NC-lazy Q
'Is the woman lazy?'

This type also includes languages in which question particles occur either in initial position or in second position, as in Ngiti (Nilo-Saharan; Democratic Republic of Congo), where there are two distinct question particles, each associated with one of these two positions (Kutsch Lojenga 1994: 379). If one position can be considered dominant, then the language is coded according to that dominant position. For example, in Abun (West Papuan; Indonesia), there is an optional sentence-initial particle *te* and an obligatory sentence-final particle *e*, as illustrated in (10).

(10) Abun (Berry and Berry 1999: 102)
(te) *nan* *nai* *nan* *bi* *suk* *it* *e*
(Q) 2SG get 2SG POSS thing COMP Q
'Have you got your things?'

Because all polar questions contain a sentence-final particle, but not all contain a sentence-initial one, Abun is treated here as a language with a sentence-final question particle.

The final type shown on the map consists of languages without question particles. This includes a number of subtypes which are not distinguished on this map but are distinguished on Map 116. It includes languages which have interrogative affixes on verbs, languages which code polar questions by a different word order, and languages in which interrogative sentences differ from declarative sentences only in intonation. Note that in some languages with question particles, it is more common not to use these particles; a language is coded here as lacking a question particle only if there is no evidence that a question particle is ever used.

In a number of languages, the position of question particles is variable. In such languages, the position of the question particle often depends on what is the focus of the question. In most if not all such languages, however, there is one position which is associated with a more neutral question, where the truth of the entire sentence is being questioned without one constituent being the focus of the question. In such cases the language is coded for this map on the basis of the position of the question particle in such neutral questions. In Imbabura Quechua (Ecuador), for example, there is an interrogative clitic that is attached after the constituent which is the focus of the question, as in (11a), where the focus of the question is *wasi-man* 'to the house'; if the sentence is a neutral question, with no specific constituent as the focus, the interrogative clitic occurs on the verb, as in (11b).

(11) Imbabura Quechua (Cole 1982: 15)
a. *wasi-man=**chu*** *ri-ju-ngui*
house-to=Q go-PROG-2
'Are you going TO THE HOUSE?'
b. *wasi-man* *ri-ju-ngui=**chu***
house-to go-PROG-2=Q
'Are you going to the house?'

Since the question particle attaches to the verb in polar questions with no constituent in focus in Imbabura Quechua, and since the normal word order in this language is verb-final, it is coded on the map as a language with a sentence-final question particle. In Syrian Arabic, which is not verb-final, the neutral position for the question particle is sentence-final (Cowell 1964: 378).

In the example in (12) from Kobon (Trans-New Guinea; Papua New Guinea), the question is signalled by an expression that involves (optionally) the word meaning 'or' plus a negated form of the verb of the sentence.

(12) Kobon (Davies 1981: 5)
ne *kaj* *ap* *mɨd-öp* ***(aka)*** ***mɨd-ag-öp***
2SG pig INDEF be-PERF.3SG **(or)** **be-NEG-PERF.3SG**
'Have you any meat?'

For the purposes of this map (and Map 116), such negative expressions are not distinguished from question particles. Such multiword expressions for forming questions are primarily found in sentence-final position.

Many languages have question particles or expressions that indicate that the question is a **leading question**, one where the speaker has an expectation as to what the answer will be. English, for example, employs what are called "tag-questions" for this purpose, illustrated in (13a), which differ from the **neutral question** form illustrated in (13b) in that their form indicates that the speaker expects a 'yes' answer.

(13) a. *You're planning to go to the beach, aren't you?*
b. *Are you planning to go to the beach?*

Similarly, Hixkaryana (Carib; Brazil) has a particle *hampe*, illustrated in (14), which signals doubt, and when used in sentences with interrogative intonation, suggests that the speaker expects the answer to be 'no'.

(14) Hixkaryana (Derbyshire 1979: 5)
ɨhme *menahno* *hampe*
egg 2>3.eat.PST DOUBT
'Did you really eat the egg?'

A neutral question in Hixkaryana, however, differs from the corresponding declarative sentence only in intonation. The map restricts attention to the expression of neutral questions, so Hixkaryana is shown on the map as only employing intonation to signal questions.

There are two ways in which a language can be coded as having a question particle at the beginning of a sentence. First, the particle may always occur at the beginning of a sentence, regardless of the word order elsewhere in the clause. Second, it may immediately precede the verb, and the dominant word order in the language may be verb-initial so that the normal position of the question particle is sentence-initial—even though if some other constituent precedes the verb, it will also precede the question particle. These two types of languages are not distinguished on the map. Similar comments apply to question particles at the end of sentences.

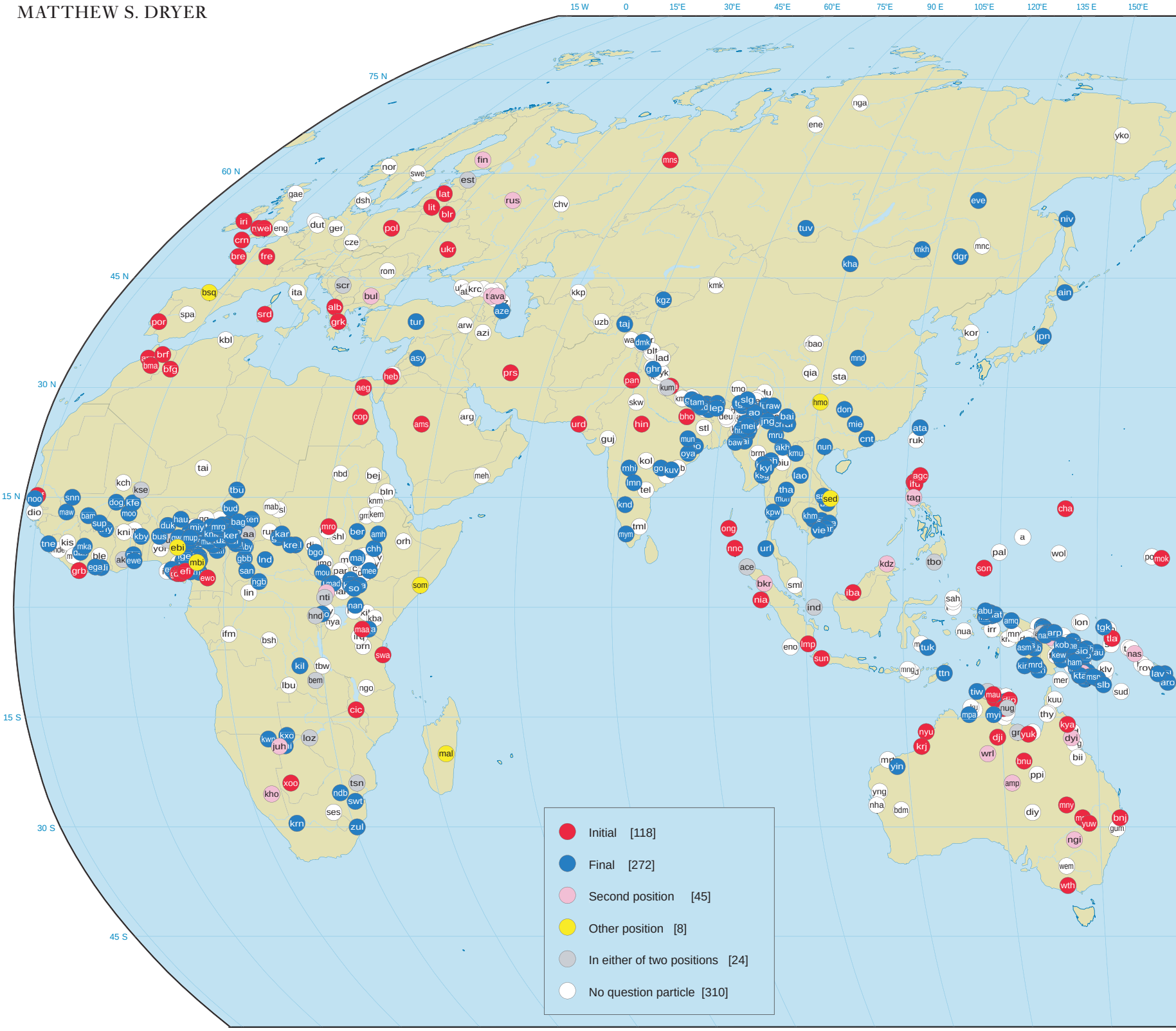
2 Geographical distribution

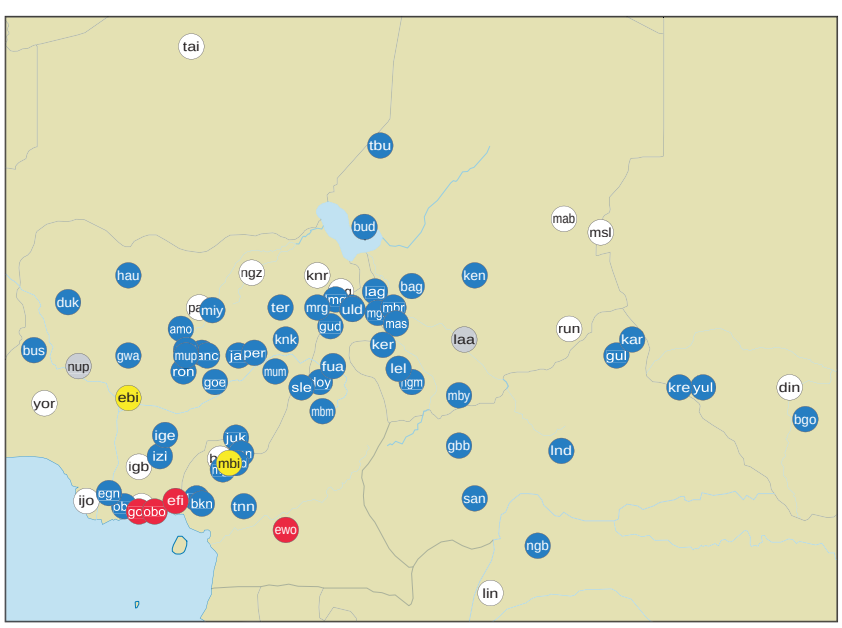
Apart from languages lacking question particles altogether, the most common type shown on the map is languages in which the question particle occurs in sentence-final position. Note, however, that the majority of languages with question particles fall into three areas: (i) a belt stretching across Africa from western Africa to central eastern Africa; (ii) an area within Asia including mainland South-East Asia and extending west into India and north through China to Japan and eastern Siberia; and (iii) New Guinea. Within these areas, sentence-final particles outnumber all other types. Outside these three areas, however, final particles are no more common than initial or second-position particles. The largest area in which final particles are absent is Europe and North Africa.

Both initial particles and second-position particles are widespread, although they are uncommon within the three areas mentioned above in which final particles predominate. Initial particles are most common in (i) Europe and North Africa; and (ii) North America, including Mesoamerica. Second-position particles are most common in North America. Languages lacking question particles are fairly evenly distributed throughout the world.

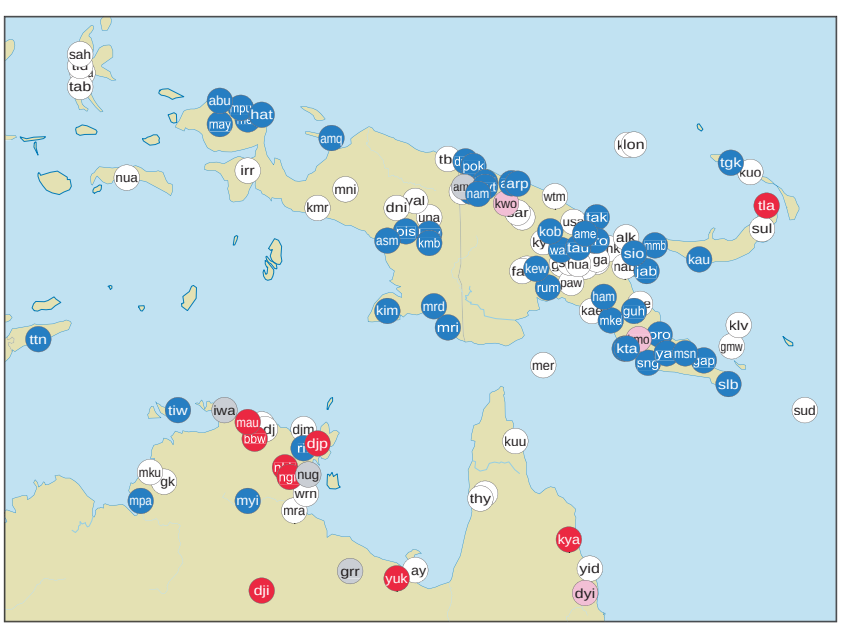
92 Position of Polar Question Particles

MATTHEW S. DRYER

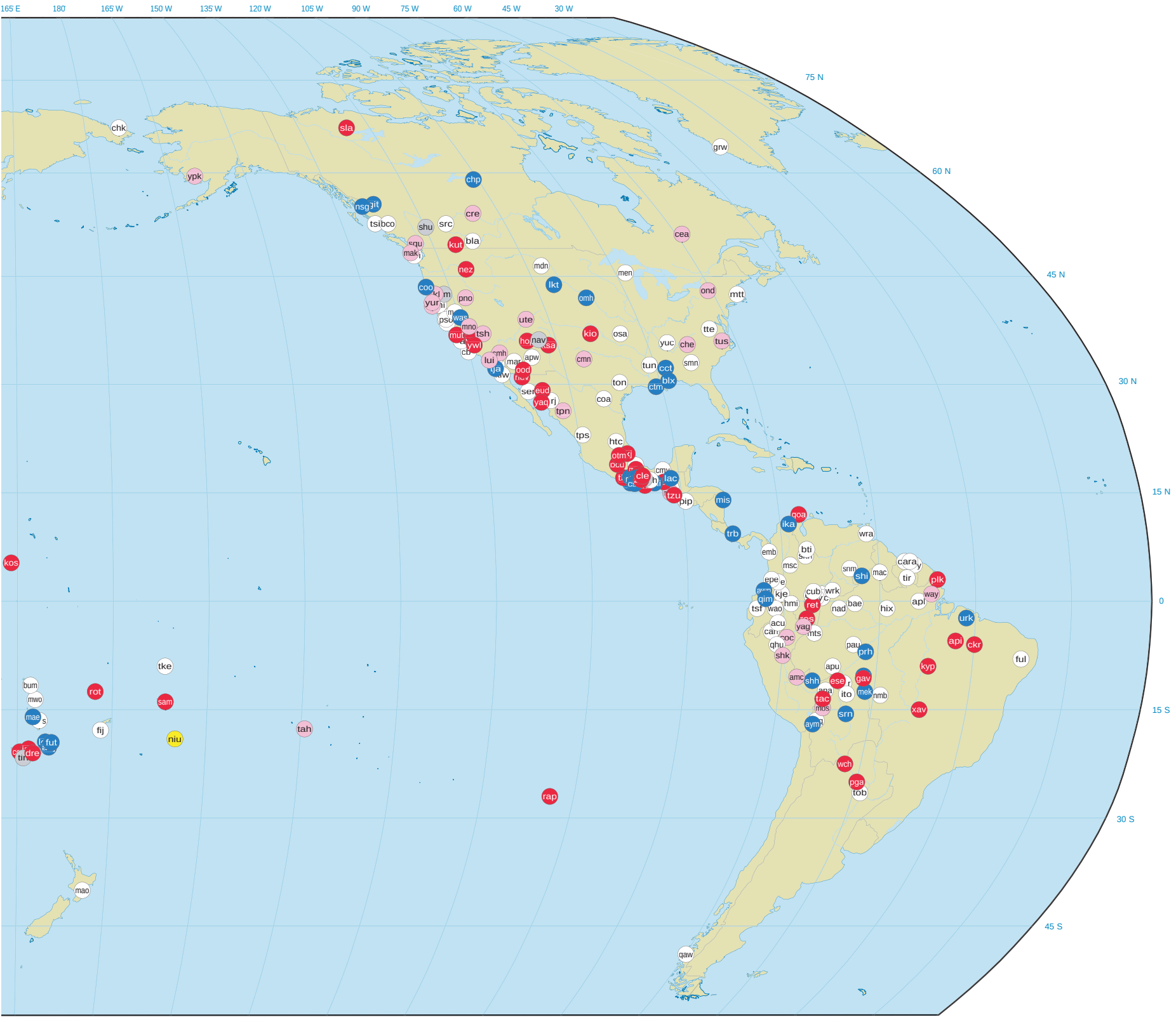


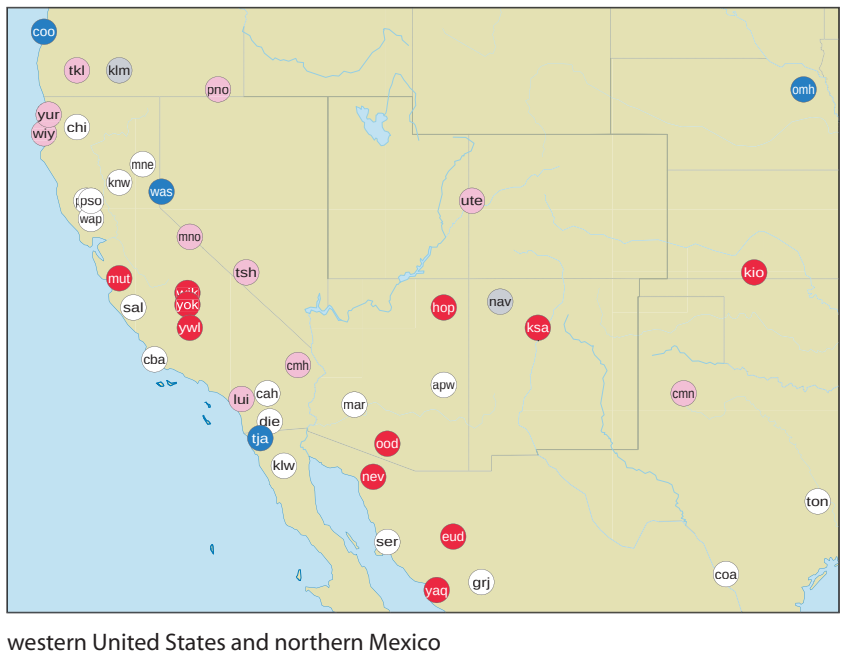


central Africa

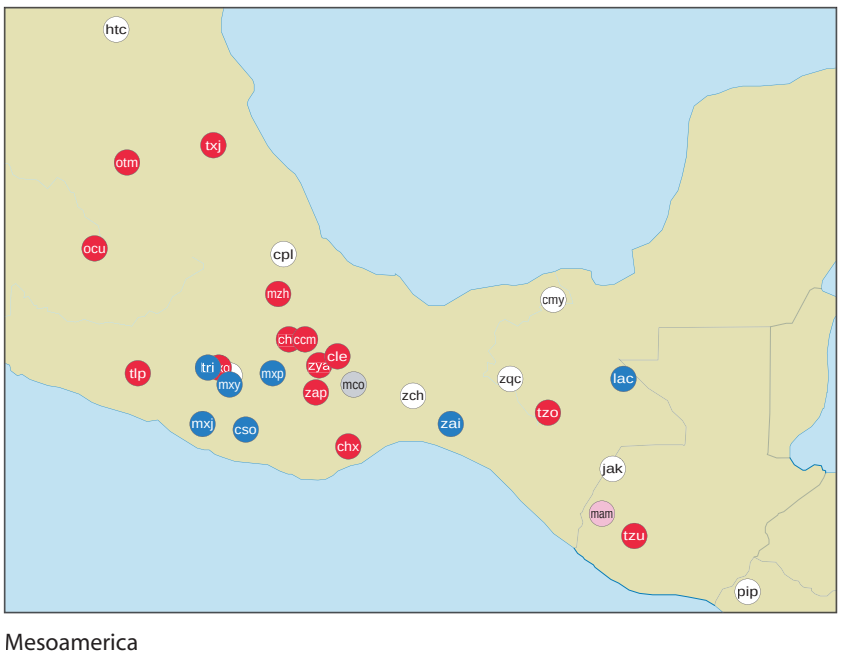


New Guinea and northern Australia





western United States and northern Mexico



Mesoamerica

93 Position of Interrogative Phrases in Content Questions

MATTHEW S. DRYER

1 Defining the values

This map shows the position of interrogative phrases in content questions. **Content questions** are questions that contain an **interrogative phrase**, like the elements in boldface in the English examples in (1).

(1) a. ***Who*** *did you see?*
b. ***Which book*** *do you want to buy?*
c. ***When*** *are you going to leave?*

Content questions differ from polar questions (see Maps 92 and 116) in that they elicit a specific answer other than 'yes' or 'no', and in containing interrogative phrases. All languages have a set of interrogative words that are characteristic of content questions, though in many languages they are identical in form to indefinite words (like *someone* in English) (see Map 46). These interrogative words typically belong to different categories (interrogative pronouns, interrogative adverbs, "interrogative adjectives", "interrogative verbs"). Because most interrogative words in English begin with *wh-*, content questions in English are often called *wh-questions*. Although interrogative phrases often consist of single words (interrogative pronouns or interrogative adverbs), they may contain additional words, as in *which book* in (1b).

1. Interrogative phrases obligatorily initial	241
2. Interrogative phrases not obligatorily initial	542
3. Mixed, some interrogative phrases obligatorily initial, some not	20
total	803

There are two common patterns cross-linguistically for the position of interrogative phrases. In some languages, like English, interrogative phrases occur **obligatorily at the beginning of the sentence**. It is possible in English to place an interrogative phrase later in the sentence, but this word order is not normally used in neutral questions. It is used either in **echo questions**, where the speaker is expressing surprise or incredulity at something that they have just heard, as in (2a), or by a teacher asking students questions, as in (2b).

(2) a. *You are leaving* ***when*** *?!?*
b. *Napoleon died in* ***what year?***

If a language allows interrogative phrases to occur in positions other than the beginning of the sentence only under special circumstances like these, it is treated as obligatorily placing interrogative phrases at the beginning of the sentence.

In other languages, interrogative phrases **do not obligatorily occur at the beginning of the sentence**, and occur naturally in other positions in the sentence, most often in whatever position is natural for the corresponding noninterrogative phrase, as in (3) from Lango (Nilotic; Uganda).

(3) Lango (Noonan 1992: 173)
òkélò *ò-nènò* ***ŋà***
Okelo 3SG-see.PERF **who**
'Who did Okelo see?'

The interrogative phrase *ŋà* 'who' in (3) occurs after the verb because Lango is an SVO language and *ŋà* is functioning as object; it thus occurs in the normal position for objects, immediately following the verb. Languages in which interrogative phrases occur in the same position as corresponding noninterrogative phrases are sometimes described as placing interrogative phrases **in situ**. There are languages, however, which obligatorily place interrogative phrases in initial position but which can nevertheless be described as placing them in situ. A number of languages normally employ a focus or cleft construction for content questions, where the focused element occurs at the beginning of the sentence. In Malagasy (Austronesian; Madagascar), for example, the normal form of a content question is that of a focus construction, which places a focused element before the verb, contrary to the normal verb-initial order in the language: (4a) illustrates a content question with the interrogative phrase in initial position, while (4b) illustrates an analogous use of the focus construction where the element in focus position is not an interrogative phrase.

(4) Malagasy (Keenan 1976: 270, 269)
a. ***amin=inona*** ***no*** *manasa* *lamba* *Rasoa*
with=what FOC wash.ACTIVE clothes Rasoa
'With what did Rasoa wash clothes?'
b. *amin=ity* *savony* *ity* *no* *manasa*
with=this soap this FOC wash.ACTIVE
lamba *Rasoa*
clothes Rasoa
'It is with this soap that Rasoa is washing clothes.'

The initial position for interrogative phrases in Malagasy thus does not involve a special construction: rather, content questions simply employ the focus construction, a reflection of the focus-like nature of interrogative phrases.

In some languages in which interrogative phrases do not occur obligatorily in sentence-initial position, they tend to occur in immediately preverbal position, as in (5) from Basque.

(5) Basque (Saltarelli et al. 1988: 6)
aita-k *Mikel* ***eraman*** *du*
father-SG.ERG Mikel **how** 3ABS.3ERG.PRES
etxera
house.SG.ALL
'How has father carried Mikel home?'

While the placement of interrogative phrases in immediately preverbal position is found in a number of verb-final languages of Europe and Asia, it appears not to be common outside of this area, though an example in Africa is Harar Oromo (Cushitic; Ethiopia; Owens 1985: 207).

A few languages exhibit at least a weak tendency to place interrogative phrases at the *end* of sentences. For example, in Tennet (Surmic, Nilo-Saharan; Sudan), the normal position for an interrogative subject in a transitive clause is after the object, contrary to the normal VSO order in the language (Randal 1998: 252–3). Noni (Broad Bantu, Niger-Congo; Cameroon) exhibits another less common position for interrogative phrases, namely immediately postverbal position (Hyman 1981: 104–5).

When we say that the position of an interrogative phrase is initial, this does not mean that the interrogative *word* is necessarily initial, since the interrogative phrase may be a larger constituent containing the interrogative word and the interrogative word may not occur at the beginning of that constituent. In most languages, for example, if the interrogative word is the object of a preposition, and if interrogative phrases occur at the beginning of the sentence, then it will be the preposition that will occur as the first word, followed by the

interrogative pronoun, as in the example in (6) from Karo Batak (Austronesian; Sumatra, Indonesia).

(6) Karo Batak (Woollams 1996: 227)
ras *isé* *kam* *ku* *jénda*
with **who** 2 to here
'Who did you come here with?'

Note that English is like this, although there is a choice (largely determined by dialect or register) between whether the preposition occurs at the beginning, with the pronoun, as in *With whom did you come here?*, and whether the pronoun occurs at the beginning of the sentence without the preposition (thereby "stranding" the preposition), as in *Who did you come here with?* English is quite unusual cross-linguistically in allowing the stranding of prepositions in content questions.

The fact that it is interrogative phrases, rather than simply interrogative words, whose position is important is also illustrated by languages in which particular interrogative modifiers of nouns follow the noun. In such languages, it will be the noun rather than the interrogative modifier that will be the first word in the sentence, as illustrated by the example in (7), again from Karo Batak.

(7) Karo Batak (Woollams 1996: 227)
arah ***apai*** *aku* *ku* *das*
way **which** 1SG to top
'Which way do I take to go up?'

Again, we can still say that the interrogative *phrase* is initial in these examples, in the sense that the interrogative phrase is the entire noun phrase containing the interrogative word.

A final though rather unusual example of an interrogative phrase containing an interrogative word that is not the first word in the sentence is also provided by Karo Batak. While the most common order in active clauses in Karo Batak is SVO, as in (8a), the normal order in active clauses in which the object is an interrogative phrase is VOS, as in (8b).

(8) Karo Batak (Woollams 1996: 214, 188)
a. *nandé* *m-bayu* *amak*
mother ACTIVE-weave mat
'Mother is weaving a mat.'
b. ***nukur*** ***kai*** *kam* *ku* *tiga*
buy **what** 2 LOC market
'What did you buy at the market?'

Since the normal position for interrogative phrases in Karo Batak is otherwise initial, as illustrated by examples (6) and (7) above, it is apparently the case in this language that the verb cannot be "stranded" without its object, just as prepositions cannot be stranded, and the interrogative phrase here is an initial phrase consisting of the verb followed by the noun phrase containing the interrogative pronoun. Similar facts obtain in Gavião (Monde, Tupian; Brazil; Moore 1984: 110).

Most languages fall fairly clearly into one or the other of the two types described above. There are two types of languages whose classification is less clear. First, in some languages, the interrogative phrase is usually placed in sentence-initial position, but this is not obligatory. Examples of languages of this sort include Mangarrayi (Merlan 1982: 6) and Awa Pit (Curnow 1997: 315). Such languages are treated here as languages in which initial position is not obligatory, unless the option of leaving an interrogative phrase in noninitial position only occurs in special circumstances, as in the English examples in (2) above. Heath (1999a: 237) describes Koyraboro Senni (Songhay; Mali) as "almost always" placing the interrogative phrase in initial position, and he describes the exceptions as being ones where placing the interrogative phrase in initial position would make the question awkward and difficult to parse. This too is treated as a special circumstance and Koyraboro Senni is shown on the map as obligatorily placing the interrogative phrase in sentence-initial position.

A second type of language which does not fall clearly into one of the first two types consists of languages where **some interrogative phrases must occur in sentence-initial position, while others need not**. For example, in Koyra Chiini (Songhay; Mali), interrogative pronouns and adverbs occur at the beginning of the sentence, in a focus construction, as in (9a) and (9b); however, noun phrases containing interrogative modifiers (one meaning 'which', another meaning 'how many') optionally remain in situ, as in (9c), and an interrogative pronoun functioning as a predicate in an equational sentence need not be initial, as in (9d).

(9) Koyra Chiini (Heath 1999b: 177, 177, 183, 178)
a. ***maa*** ***na*** *wor* *o* *fai* *hondu* *jiiroo*
what FOC 2PL IMPF grow dune this.year
'What will you (pl) grow on the dune this year?'
b. *man* *na* *ni* *hun*
where FOC 2SG leave
'Where did you come from?'
c. *afoo* *kul* *go* *hima* *ka* *jow* ***ije*** ***merje***
one all IMPF should INF take **piece** **how.many**
'How many pieces should each one take?'
d. *woo* *či* ***mey***
that be **who**
'Who is that?'

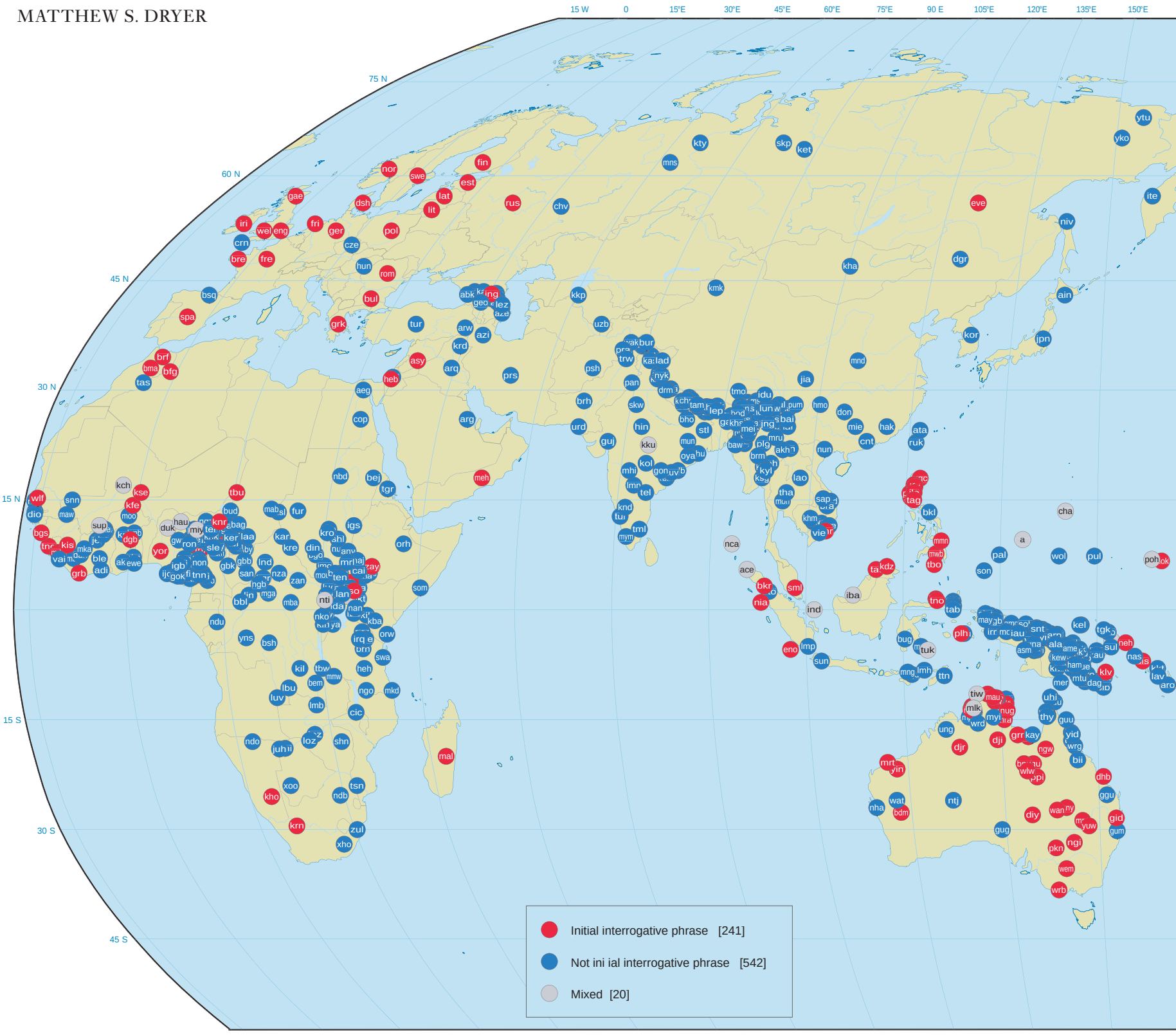
In Malakmalak (Daly; Northern Territory, Australia), interrogative noun phrases occur in initial position but interrogative adverbs need not (Birk 1976: 26). In Tukang Besi (Austronesian; Sulawesi, Indonesia), interrogative noun phrases functioning in predicates must be initial while other interrogative noun phrases need not be (Donohue 1999: 355, 451). And in Pohnpeian (Oceanic; Micronesia), the interrogative word for 'why' must be initial; other interrogative phrases do not have to be (Rehg 1981: 314–23).

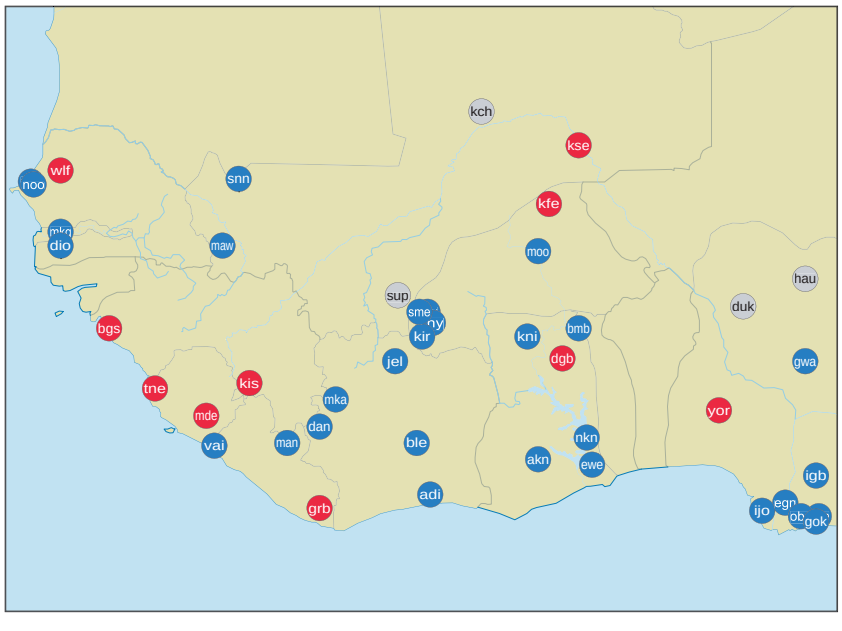
2 Geographical distribution

The two major types represented on the map are both widespread, but there are clear geographical patterns. Areas in which languages predominate which obligatorily place interrogative phrases in initial position include (i) Europe and North Africa; (ii) the Philippines and the adjacent areas; (iii) much of Australia, except the northern part of Northern Territory and Cape York in the north-east; and (iv) the Americas. Areas in which languages predominate which do not obligatorily place interrogative phrases in initial position include (i) much of Africa, though less so in the west and north-west; (ii) most of the mainland of Asia, plus Japan; and (iii) New Guinea and the Pacific. Languages of a mixed type seem also to exhibit some geographical clustering: (i) there are a number in the general vicinity of Indonesia and the western Pacific; and (ii) there are also a number in West Africa.

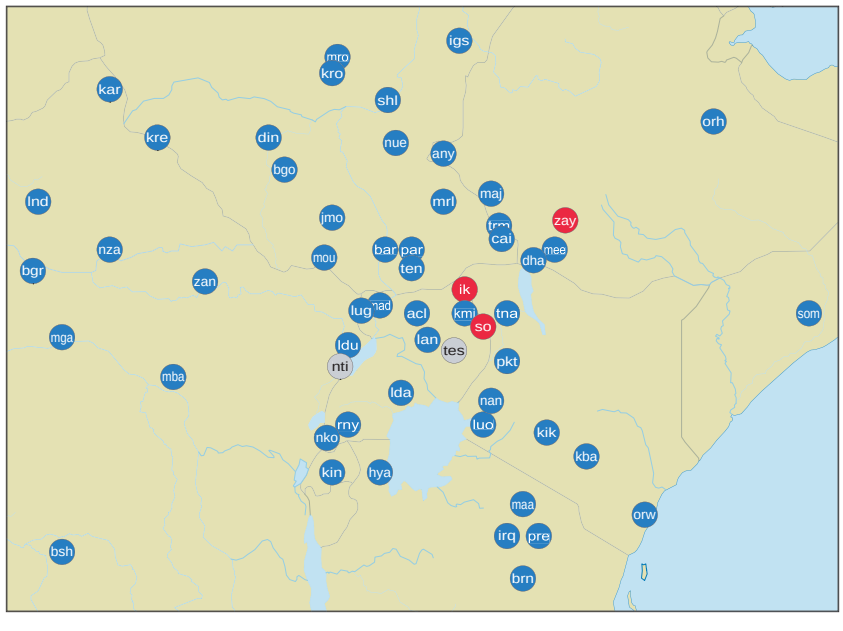
93 Position of Interrogative Phrases in Content Questions

MATTHEW S. DRYER

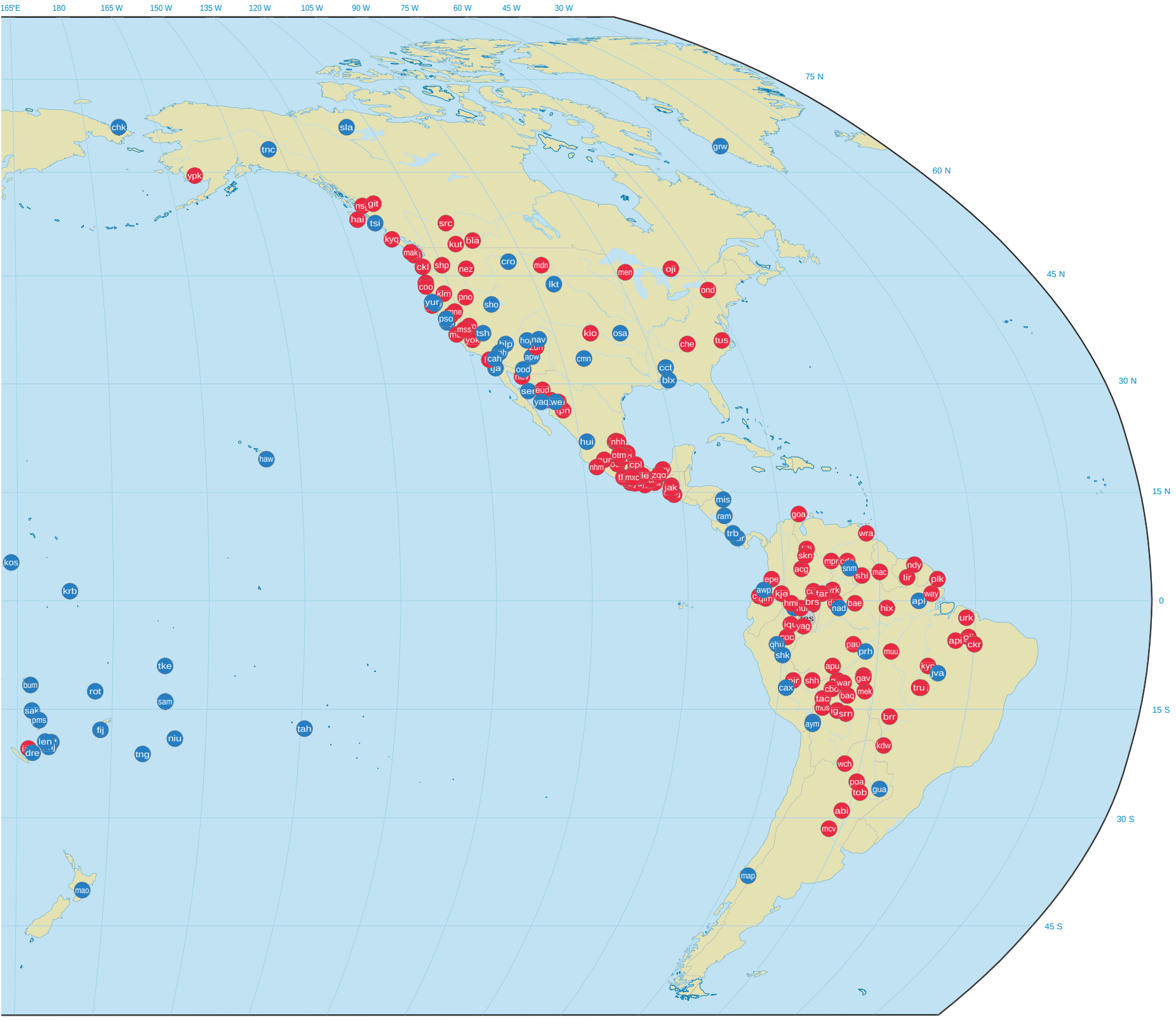


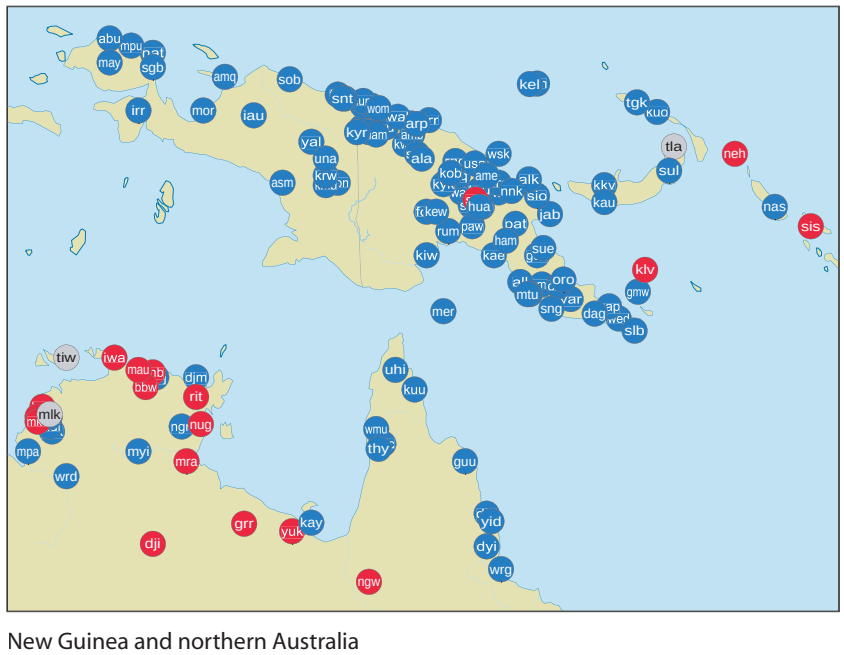


West Africa

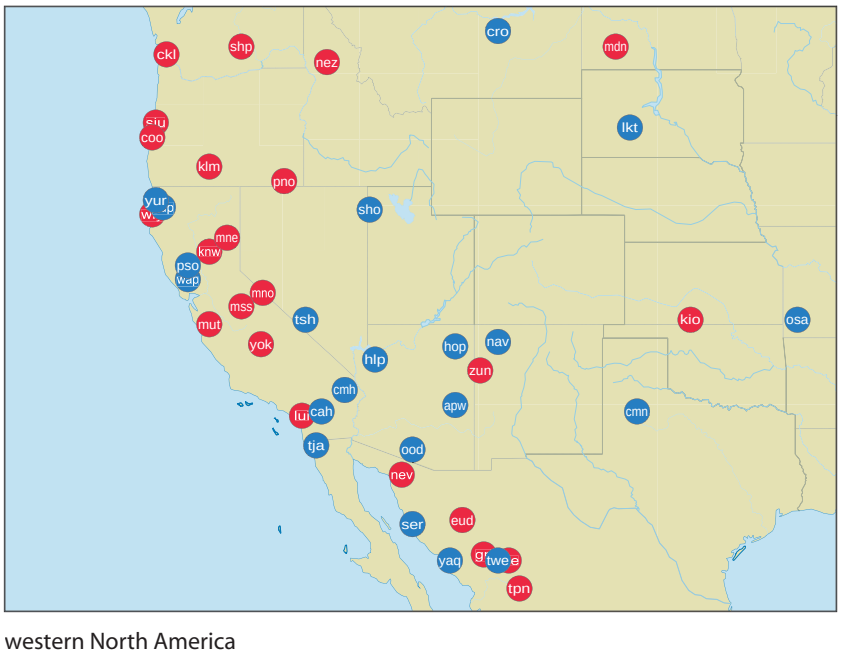


eastern Africa





New Guinea and northern Australia



western North America

94 Order of Adverbial Subordinator and Clause

MATTHEW S. DRYER

1 Defining the values

This map shows the position of adverbial subordinators, morphemes which mark adverbial clauses for their semantic relationship to the main clause. Adverbial subordinators which are separate words are referred to as *subordinating conjunctions* by traditional grammar. Examples of adverbial subordinators in English include *because, although, when, while*, and *if*. The map further distinguishes adverbial subordinators which are separate words from ones which are affixes on verbs.

1. Adverbial subordinators which are separate words and which appear at the beginning of the subordinate clause	367
2. Adverbial subordinators which are separate words and which appear at the end of the subordinate clause	90
3. Clause-internal adverbial subordinators	8
4. Suffixal adverbial subordinators	59
5. More than one type of adverbial subordinator with none dominant	87
total	611

The first type shown on the map includes languages in which the **adverbial subordinators are separate words that appear at the beginning of the clause**, as in the English example in (1a) and the example in (1b) from Hanis Coos (Coosan; Oregon).

(1) a. ***because*** *it was Friday*

b. Hanis Coos (Frachtenberg 1913: 72)
yanL *tsîʹx·tī* *eᵉ-heʹlaq*
if there 2SG-arrive
'if you arrive there'

The second type shown involves languages in which **adverbial subordinators are separate words that appear at the end of the clause**, as in the example in (2a) from Kombai (Trans-New Guinea; Papua, Indonesia) and the example in (2b) from Kolyma Yukaghir (isolate; Siberia, Russia).

(2) a. Kombai (de Vries 1993: 58)
khe-khino *rerakharu* ***rofode***
his-legs swollen **because**
'because his legs are swollen'

b. Kolyma Yukaghir (Maslova 2003: 399)
ulum *gud-uj-l'ie-t* ***tit***
mad become-ITER-INGR-SS.IMPF **although**
'although he was going mad'

The third type shown involves languages in which **the adverbial subordinator appears inside the clause**. In Nkore-Kiga (Bantu; Uganda), the adverbial subordinator appears between the subject and the verb, as in (3).

(3) Nkore-Kiga (Taylor1985: 27)
wa-kami ***obu*** *y-aa-tuuriza* *enjojo*
Mr.-rabbit **when** he-TODAY.PST-challenge elephant
'when Brer Rabbit challenged the elephant'

In Yukulta (Tangkic; Queensland, Australia), the subordinator is an enclitic, attaching to the first constituent of the clause, as in (4), where *=aka* attaches to the first word in the subordinate clause, *ŋapay*, and precedes the present-tense clitic *=ŋka*.

(4) Yukulta (Keen 1983: 243)
pukawaṭa-ṇiŋki, *[ŋapay=**aka**=ŋka* *kalka]*
die.IND-he.FUT [very=**because**=PRES sick]
'He might die, because he's very sick.'

There are only eight languages shown on the map with this as the dominant type, but there are many other languages in which this type occurs, though not as the dominant type of subordinator.

The fourth type shown involves **languages in which the adverbial subordinator is a suffix on the verb**, illustrated in (5a) for Kiowa (Kiowa-Tanoan; central United States) and in (5b) for Hunzib (Daghestanian; eastern Caucasus, Russia).

(5) a. Kiowa (Watkins 1984: 242)
*à-dę̀·kʼɔ́·**-àl*** *hɔ́n* *àn* *à-dę̀·hɛ́·m-ô*
1SG-lie-**although** NEG HAB 1SG-sleep-NEG
'Although I lie down, I can't fall asleep.'

b. Hunzib (van den Berg 1995: 98)
zaƛe *n-ex-**áyd**,* *xõx-ƛʼo* *ƛibu* *zuqʼu-r*
wind NC5-strike-**before** tree-SUP leaf be-PRET
'Before the wind blew, there were leaves on the tree.'

In most languages of this sort, the language is verb-final and the suffixes on the verb are also at the end of the clause. A logically possible type for which I have no clear example is a language where the adverbial subordinator is a prefix on the verb.

The final type shown involves **languages in which more than one of the above types occurs** but in which no single type is dominant. This includes languages with both affixes and separate words functioning as adverbial subordinators, languages with both clause-initial and clause-final words as adverbial subordinators, and languages combining clause-internal or second-position adverbial subordinators with clause-initial or clause-final ones, without one being dominant. For example, Tigre (Semitic; Eritrea) has both initial subordinators, as in (6a), and internal subordinators, as in (6b).

(6) Tigre (Raz 1983: 91)

a. ***ḥaqo*** *ʔət* *karan* *ʔaqabbəl*
after to Karan 1SG.return
'after I return to Karan'

b. *šimat* ***ʔət*** *laḥazzu*
chieftainship **while** they.seek
'while they were seeking the chieftainship [for him]'

Bwe Karen (Tibeto-Burman; Myanmar) has initial subordinators, as in (7a), final subordinators, as in (7b), internal subordinators, as in (7c), and combinations of internal and final ones, as in (7d).

(7) Bwe Karen (Henderson 1997: 395, 23, 230, 78)

a. *yə-khɔ́* *ge* ***[kɔ́*** *yə-dɛ-mɛ*
1SG-FUT go.back [**when** 1SG-NMLZ-do
wá *lɔ]*
complete DECL]
'I shall go back when my work is done.'

b. *gəli* *ə-co* *u* ***ɓaʃɔ́*** *nu*
wind NMLZ-cool blow although PCL
'even though there was a cool breeze blowing'

c. *kə-a* ***má**-cɛ,* . . .
1PL-eat **if**-3.OBJ
'if we ate it, . . .'

d. *nə-**ɗé*** *ɔ* ***khalɛ́***
2SG-**if** be.at **if**
'if you stay'

In Epena Pedee (Choco; Colombia), some of the subordinators are clause-final words, as in (8a), while others are verbal suffixes, as in (8b).

(8) Epena Pedee (Harms 1994: 162, 161)

a. *pɨ́* *mimia-rú* ***pérã***
you work-PRES **because**
'because you work'

b. *hára* *pʰã-dá-**irã***
tell be.PL-PL-**because**
'because they tell me'

The suffix *-irã* 'because' is used on auxiliary verbs, while the separate word *pérã* 'because' is used when there is no auxiliary. In Majang (Surmic; Ethiopia), adverbial clauses tend to involve a word at the beginning of the clause and a clitic at the end. With 'because' clauses there is also a suffix on the verb, so the meaning 'because' is coded by three simultaneously occurring morphemes, as in (9).

(9) Majang (Unseth 1989: 117)
agutucee=ko *tolay* *ɗoko-**ɗu*** *ogol=**ku***
because=PST Tolay bring-**reason** mead=**reason**
'because Tolay brought mead'

In some languages with more than one form of subordinator, one can be treated as dominant (see the box section "Determining Dominant Word Order" on p. 371). For example, in Miya (Chadic; Nigeria), there are both clause-initial subordinators, as in (10a), and subordinators which appear in preverbal position, as in (10b) and (10c). The latter sort follow the subject pronoun if there is one, as in (10b); but if the subject is nominal, the nominal subject will follow the verb in such clauses, with the result that the subordinator is initial, as in (10c). Although initial in (10c), its position might be more accurately described as immediately preverbal, with the understanding that in clauses of this sort, immediately preverbal position will also happen to be initial position in the clause. Because some subordinators are always clause-initial (10a) while others are only sometimes clause-initial (10b,c), Miya is shown on the map as having initial subordinators.

(10) Miya (Schuh 1998: 370, 371)

a. . . . ***àadama*** *fa* *ta* *bíy* *kərá=ɯ*
so.that you NEG PCL steal=NEG
'. . . so that you would not steal it'

b. *mən* ***kwaa*** *zara-tlən* *ka,* . . .
1SG **when** call-3PL DEF
'When I call them, . . .'

c. ***kwáa*** *zàra-tlən* *aa* *Kàsham* *ká,* . . .
when call-3PL SUBJ Kasham DEF
'When Kasham calls them, . . .'

Particularly when coding languages with adverbial subordinator affixes, care was taken not to include affixes which are general markers of subordination, or affixes which may be more properly viewed as part of the tense-aspect system. For example, an affix on a subordinate verb indicating that the event of the subordinate clause is simultaneous with that of the main clause has a meaning somewhat analogous to that of *while* in English, but was considered a tense-aspect affix rather than an adverbial subordinator for the purposes of this map.

In some languages, the words treated here as adverbial subordinators are formally adpositions combining with nominalized forms of verbs. For example, in (11) from Krongo (Kadugli; Sudan), the word *kárrí* 'after' is formally a preposition combining with a nominalized form of the verb in genitive case, the case that this preposition governs.

(11) Krongo (Reh 1985: 342)
kárrí *má-t-ÙUy* *sháày*
after GEN-NMLZ-drink tea
'after drinking tea'

With nouns, this preposition means 'behind'.

Similarly included among suffixal adverbial subordinators on the map are morphemes that are formally case suffixes. For example, in Kannada (Dravidian; India), the instrumental case suffix is used with gerund clauses to form 'because' clauses, as in (12).

(12) Kannada (Sridhar 1990: 74)
bisiḷu *hecca:giruvudar-**inda***
heat much.ADV.be.N.PST.GERUND.OBL-INSTR
'since it's very hot'

2 Geographical distribution

The most frequent type shown is adverbial subordinators which are separate words occurring at the beginning of the clause. This type is represented by a majority of the languages on the map. It is especially common in (i) Europe; (ii) an area in Asia stretching from the Middle East to India; (iii) South-East Asia through the Pacific, including Australia but not the mainland of New Guinea; (iv) Africa; and (v) parts of North America, notably the Pacific North-West and Mesoamerica.

Clause-final subordinators that are separate words are common in (i) an area in Asia stretching from India north-east through Myanmar and China into northeastern Asia; (ii) New Guinea; and (iii) South America. Suffixal adverbial subordinators are generally found in the same areas as clause-final subordinators that are separate words, but are also common in the Caucasus region and in the United States.

Languages with a mixture of types are widely distributed, but less common in areas where clause-initial subordinating words predominate, except across central Africa.

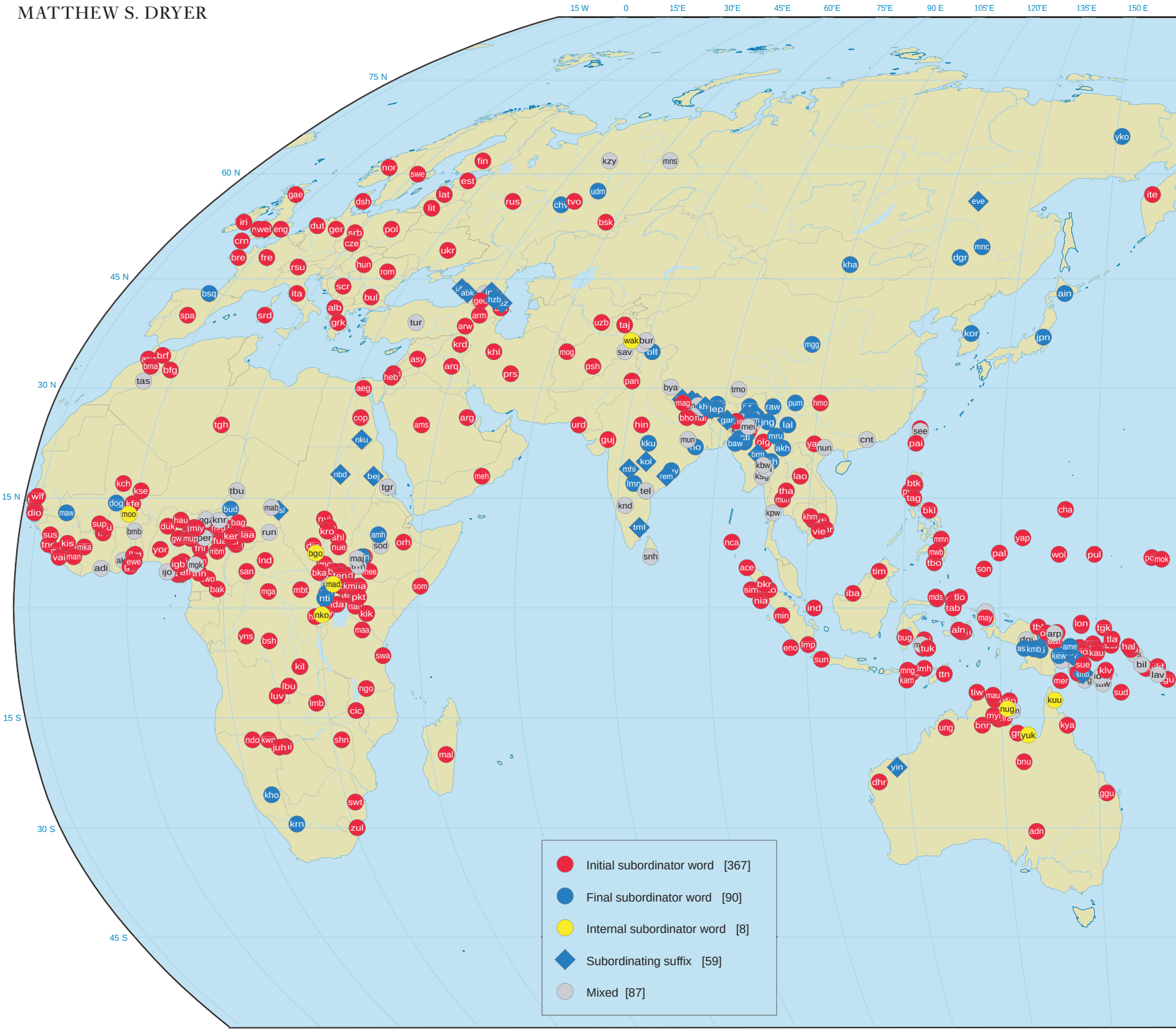
3 Theoretical issues

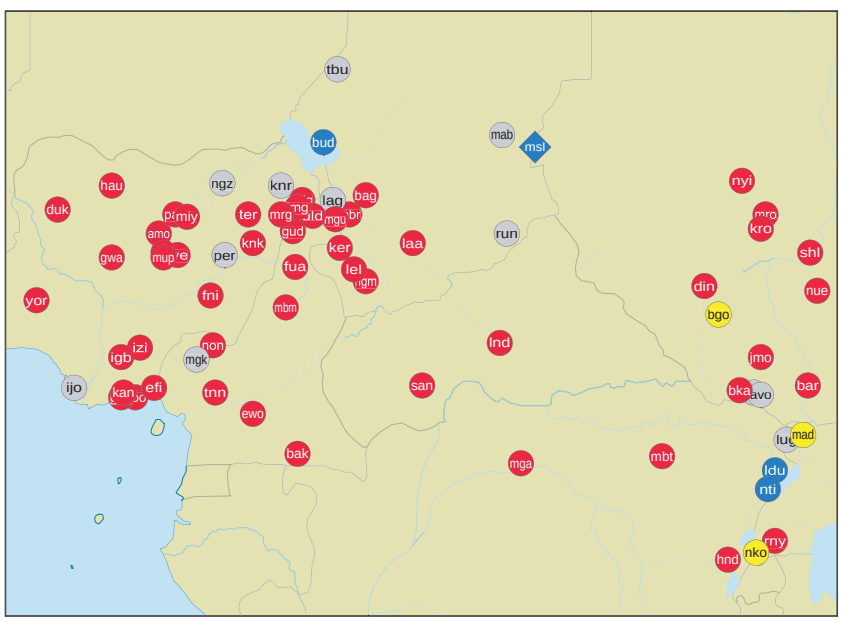
The map displays two interesting asymmetries of word order and affix position. First, among adverbial subordinators that are separate words, clause-initial subordinators outnumber clause-final subordinators by over three to one. If we add to the latter type verbal suffixes, which most commonly occur in clause-final position in verb-final languages, the initial subordinators are still more than twice as common. The second asymmetry is among affixal adverbial subordinators, all clear instances of which are suffixes, with no clear instances of prefixes. While this fits into the overall preference for suffixes (see Map 26), it constitutes one of the stronger instances of this preference. Adverbial subordinators resemble adpositions and case affixes in that they both indicate the semantic relationship between a constituent and a verb, so the rarity of prefixal adverbial subordinators can be viewed as related to the rarity of case prefixes, although why both of these are rare is not clear.

For a detailed study of adverbial subordinators among the languages of Europe, see Kortmann (1997). See Diessel (2001) for discussion of a related typological feature, the order of adverbial clause and main clause.

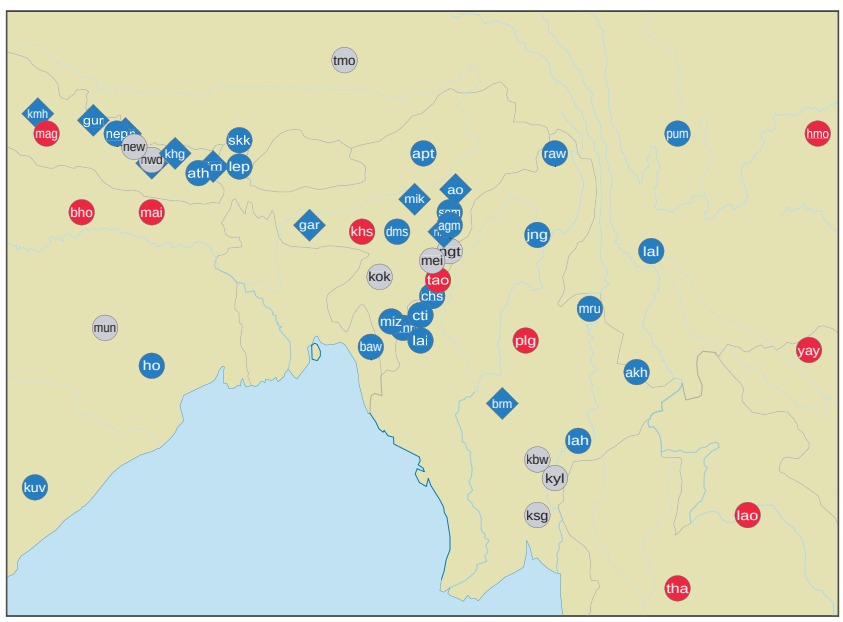
94 Order of Adverbial Subordinator and Clause

MATTHEW S. DRYER

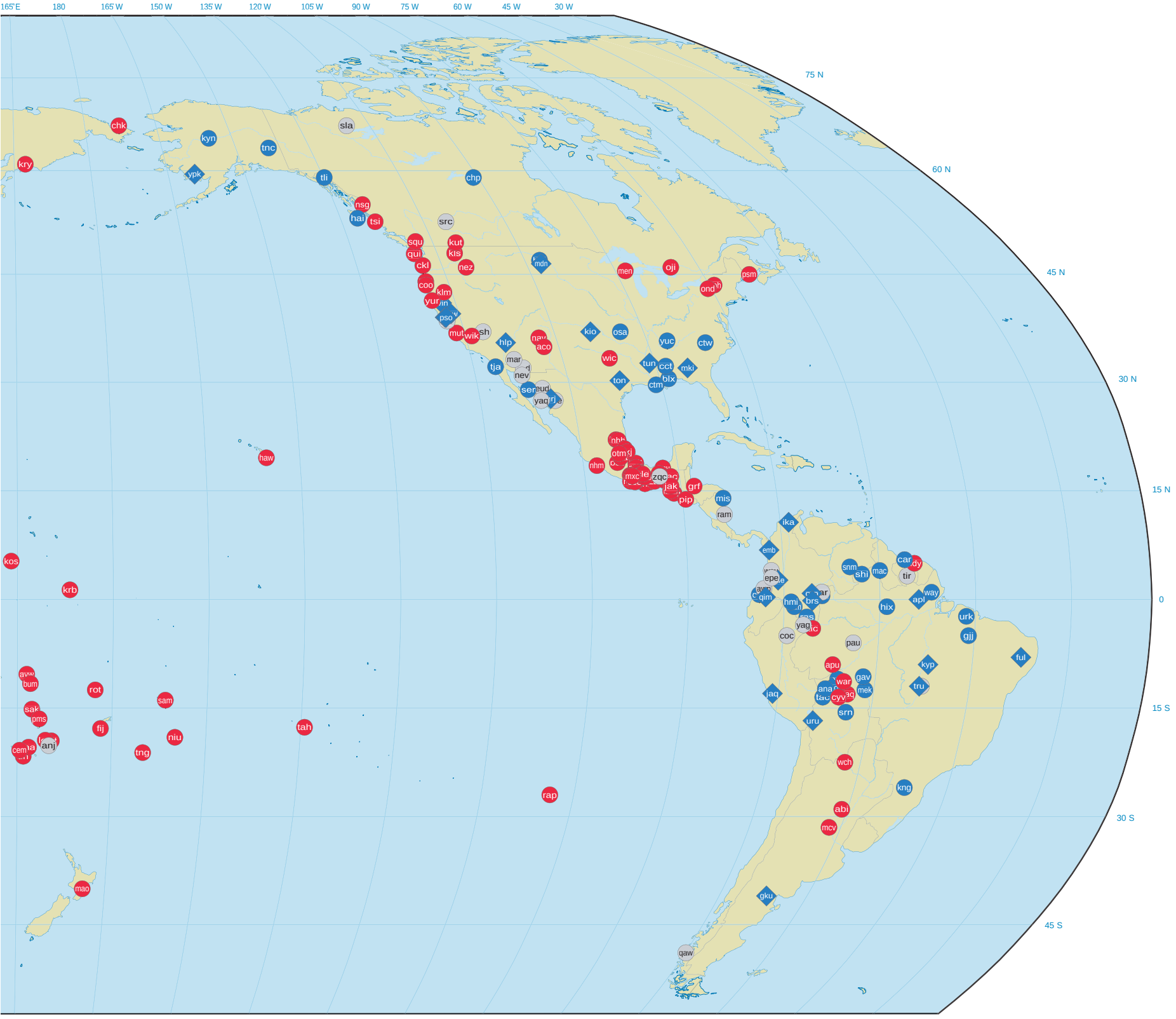


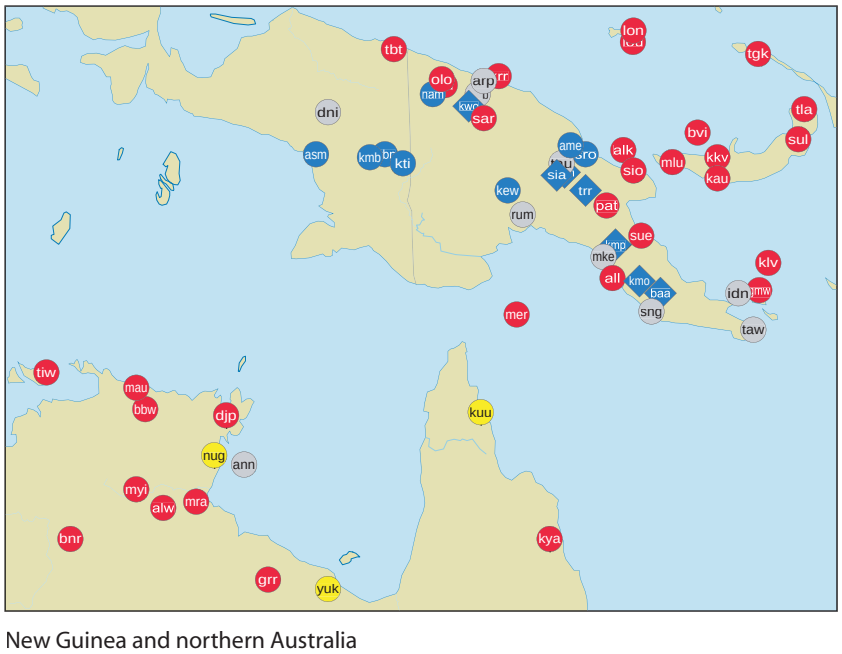


central Africa

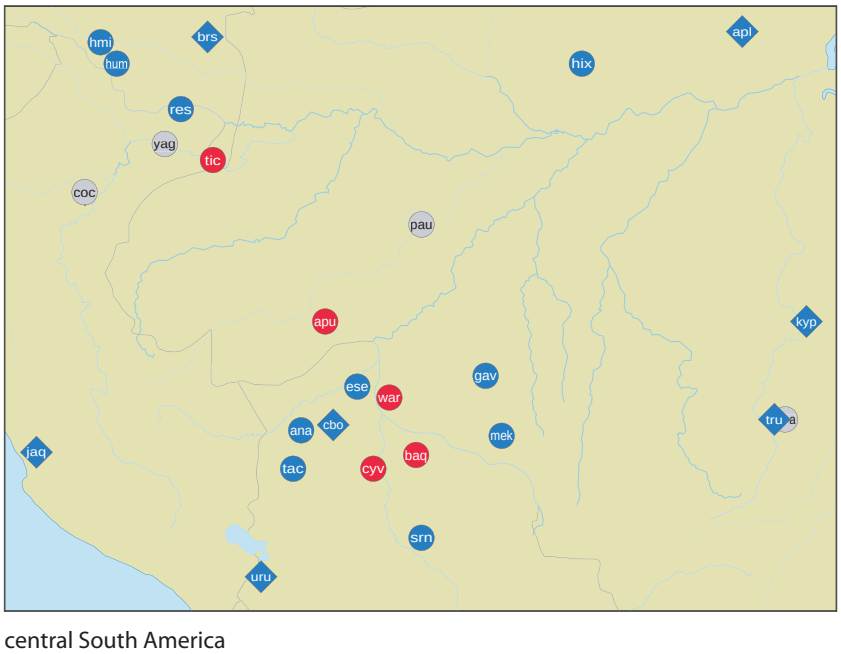


from Nepal to Thailand





New Guinea and northern Australia



central South America

95 Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase

MATTHEW S. DRYER

1 Defining the values

One of the central activities in linguistic typology is the investigation of relationships between different typological features. Maps 95, 96, and 97 illustrate three sorts of relationship between typological features. If one takes two typological features with two values each, they intersect to define four possible language types. This map, Map 95, illustrates a case where two of these types are very common and two are relatively uncommon. Map 96 illustrates a case where three of these types are very common and one is relatively uncommon. And Map 97 illustrates a case where there is no interaction between the two features, where all four types are very common. Each of these three maps shows the relationship between two features that are shown separately on other maps in this atlas.

This map shows the relationship between two of the values shown in Map 83, OV (object–verb order) and VO (verb–object order), and two of the values shown on Map 85, postpositions versus prepositions. The intersection of these defines the first four types shown in the box. Two of these four types, the first and fourth, are very common, while the other two types, the second and third, are relatively uncommon.

1. Object–verb and postpositional (OV&Postp)	427
2. Object–verb and prepositional (OV&Prep)	10
3. Verb–object and postpositional (VO&Postp)	38
4. Verb–object and prepositional (VO&Prep)	417
5. Languages not falling into one of the preceding four types	141
total	1033

The example in (1) from Nivkh (isolate; Sakhalin Island, Russia) illustrates the first type, with **OV order and postpositions**; (1a) illustrates the OV order while (1b) illustrates a postposition *rulku* 'across'.

(1) Nivkh (Gruzdeva 1998: 18, 38)

a. *n'iɣvŋ liɣř k'u-nt*
man wolf kill-FIN
'The man killed a wolf.'

b. *n'yŋ pila eri rulku vi-d'*
1PL big river across go-FIN
'We went across a big river.'

This type is very common, representing the typical situation found in OV languages. The examples in (2a) from Persian and in (2b) from Tigre (Semitic; Eritrea) illustrate the second type, **OV order with prepositions**.

(2) a. Persian (Mahootian 1997: 139)
pul-o be ma dad
money-OBJ to 1PL gave
'He gave the money to us.'

b. Tigre (Raz 1983: 108)
ʔana ʔət lehay māy ʔət-ta maḥāz gale
1SG in that water in-the river something
ʔərʔə halle-ko
see be-1SG
'I see something in the water in the river.'

This type is relatively uncommon; it is atypical for an OV language to have prepositions. The third type, **VO with postpositions**, is also uncommon; it is atypical for VO languages to have postpositions. The examples in (3) from Koyra Chiini (Songhay; Mali) illustrate this type, (3a) illustrating the VO order and (3b) a postposition.

(3) Koyra Chiini (Heath 1999b: 103, 124)

a. *har di o guna woy di*
man DEF IMPF see woman DEF
'The man sees the woman.'

b. *ay too huu di ra*
1SG arrive house DEF LOC
'I arrived in the house.'

This type is also illustrated in (4), from Arawak (Arawakan; Suriname), (4a) illustrating the VO order, (4b) a postposition.

(4) Arawak (Pet 1987: 47, 83)

a. *li fary-fa aba kabadaro*
he kill-FUT one jaguar
'He will kill a jaguar.'

b. *da-dykha no hyala diako*
1SG-see it bench on
'I saw it on a bench.'

The example in (5), from Chalcatongo Mixtec (Oto-Manguean; Mexico), illustrates the fourth type, VO order and prepositions.

(5) Chalcatongo Mixtec (Macaulay 1996: 117)
ni-sáʔa=rí ɨ̃ɨ̃ šiò xakúu sesɨ́ʔí=ri
COMP-make=1.SUBJ one skirt for daughter=1.POSS
'I made a skirt for my daughter.'

This type is relatively common, the typical situation in VO languages. English is another example of a language of this type.

The fifth type shown on the map is for languages not falling into one of the first four types. Most of these are languages lacking a dominant order of object and verb or a dominant order of adposition and noun phrase. It also includes languages with inpositions and languages lacking adpositions altogether.

2 Geographical distribution

The distribution of the two common types on the map, OV with postpositions and VO with prepositions, is similar to the distribution of OV and VO order on Map 83 and the distribution of postpositions and prepositions on Map 85. OV order with postpositions is especially common in (i) Asia, except in the Middle East and South-East Asia; (ii) New Guinea; and (iii) North America, except in the Pacific North-West and in Mesoamerica. It is also found as the more frequent type in (i) Australia; (ii) South America; and (iii) two areas in Africa, one around Ethiopia, Sudan, and eastern Chad, the other in West Africa around Mali, Burkina Faso, western Côte d'Ivoire, and Liberia. VO order with prepositions is especially common in (i) Europe; (ii) sub-Saharan Africa, except in the two areas noted above; (iii) an area stretching from South-East Asia eastward into the Pacific; (iv) the Pacific North-West; and (v) Mesoamerica.

Apart from sporadic instances, VO order with postpositions is found in (i) an area in West Africa including Benin, Togo, Ghana, and eastern Côte d'Ivoire; (ii) an area in northern Europe centred in Finland; and (iii) scattered languages throughout South America. In fact, among the VO languages of South America, postpositions are as common as prepositions. OV languages with prepositions, the least frequent type, are only found sporadically. The only area

where the map shows repeated instances in the same area is among Iranian languages in Iraq, Iran, and Tajikistan.

Many instances of languages of the two less common types can be seen as being located near or between languages of both of the two more common types. For example, the VO and postpositional languages in West Africa have OV and postpositional languages to the west and north-west, and VO and prepositional languages to the east. The OV and prepositional languages in and around Iran have OV and postpositional languages to the north and east, and VO and prepositional languages to the south and west. In other instances, this is not the case: for example, there are a couple of VO postpositional languages in northeastern Peru with no VO prepositional languages nearby.

3 Theoretical issues

The fact that only two of the four types shown on the map are common shows that these two typological features correlate. The distribution can be described by a **bidirectional implicational universal**: a language tends to be OV if and only if it is postpositional. In other words, if a language is OV then it is usually postpositional, and if a language is postpositional then it is usually OV. This bidirectional nature contrasts with the pattern illustrated on Map 96, where the implication is not bidirectional.

The correlation between the order of object and verb and the order of adposition and noun phrase is only one of many correlations between the order of object and verb and the order of other pairs of elements, and there is considerable literature discussing these correlations and attempting to explain them (including Greenberg 1963, Hawkins 1983, 1994, Dryer 1992). Among the various hypotheses that have been put forward are (i) the idea that using prepositions in VO languages and postpositions in OV languages involves consistent ordering of grammatical heads with respect to their dependents or complements; (ii) the idea that the word-order correlations reflect a tendency to use consistently left-branching structures or consistently right-branching structures, ultimately motivated by considerations of ease of processing; and (iii) the idea that the correlations reflect the results of grammaticalization: adpositions arise from verbs, retaining their order with respect to the noun phrase, or from head nouns in genitive constructions. Significantly, of the twenty-three VO languages with postpositions for which I have data on the genitive construction, twenty employ genitive–noun order, and two allow both orders of genitive and noun with neither order dominant; only one VO postpositional language has noun–genitive order. For example, Koyra Chiini, cited in (3) above illustrating a VO language with postpositions, has genitive–noun order, as in (6).

(6) Koyra Chiini (Heath 1999b: 85)
[ay baaba wane] huu di
[1SG father POSS] house DEF
'my father's house'

This fits the prediction of grammaticalization theory that an adposition grammaticalizing from a head noun in a genitive construction in a VO language with genitive–noun order will become a postposition, although it equally well could simply be that VO postpositional languages are former OV postpositional languages. The reverse situation obtains with OV languages that have prepositions: of the nine languages of this type for which I have data on their genitive construction, seven employ noun–genitive order, and two have both orders with neither dominant; none of them have genitive–noun order as the dominant order, even though OV languages generally place the genitive before the noun. For example, Tigre, illustrated in (2b) as an example of an OV language with prepositions, places the genitive after the noun, as in (7).

(7) Tigre (Raz 1983: 108)
walat farʕon
daughter Pharaoh
'the daughter of the Pharaoh'

There are a number of other observations that can be made about the two less common types shown on this map. First, although both are uncommon, one of them, VO with postpositions, is somewhat more common than the other (though the difference may not be significant): the map shows thirty-nine instances of VO languages with postpositions, but only eleven instances of OV languages with prepositions. Second, when one examines the languages of the uncommon types in greater detail, one finds that they are often not typical OV or VO languages. This observation is illustrated by the number of languages mentioned above where the order of genitive and noun is unexpected given the order of object and verb in the language. Similarly, among the ten OV languages with prepositions, two are OVS, a quite unusual word order (see Map 81). Also, six of them place adpositional phrases after the verb either with as great a frequency as in preverbal position or at least as a possible option; OV languages typically place adpositional phrases before the verb (see Chapter 84 and Dryer 1992: 92). For example, Iraqw (Cushitic; Tanzania) employs OV order, as in (8a), but prepositional phrases follow the verb as often as they precede it, as in (8b).

(8) Iraqw (Mous 1993: 238, 226)
a. *ta tiʔitá-r axwees*
IMPERS.SUBJ story-F tell.3SG.M
'They tell a story.'
b. *baabá aa búuhh as aníng*
father 3.SUBJ.PERF be.angry.3SG because.of 1SG
'Father is angry because of me.'

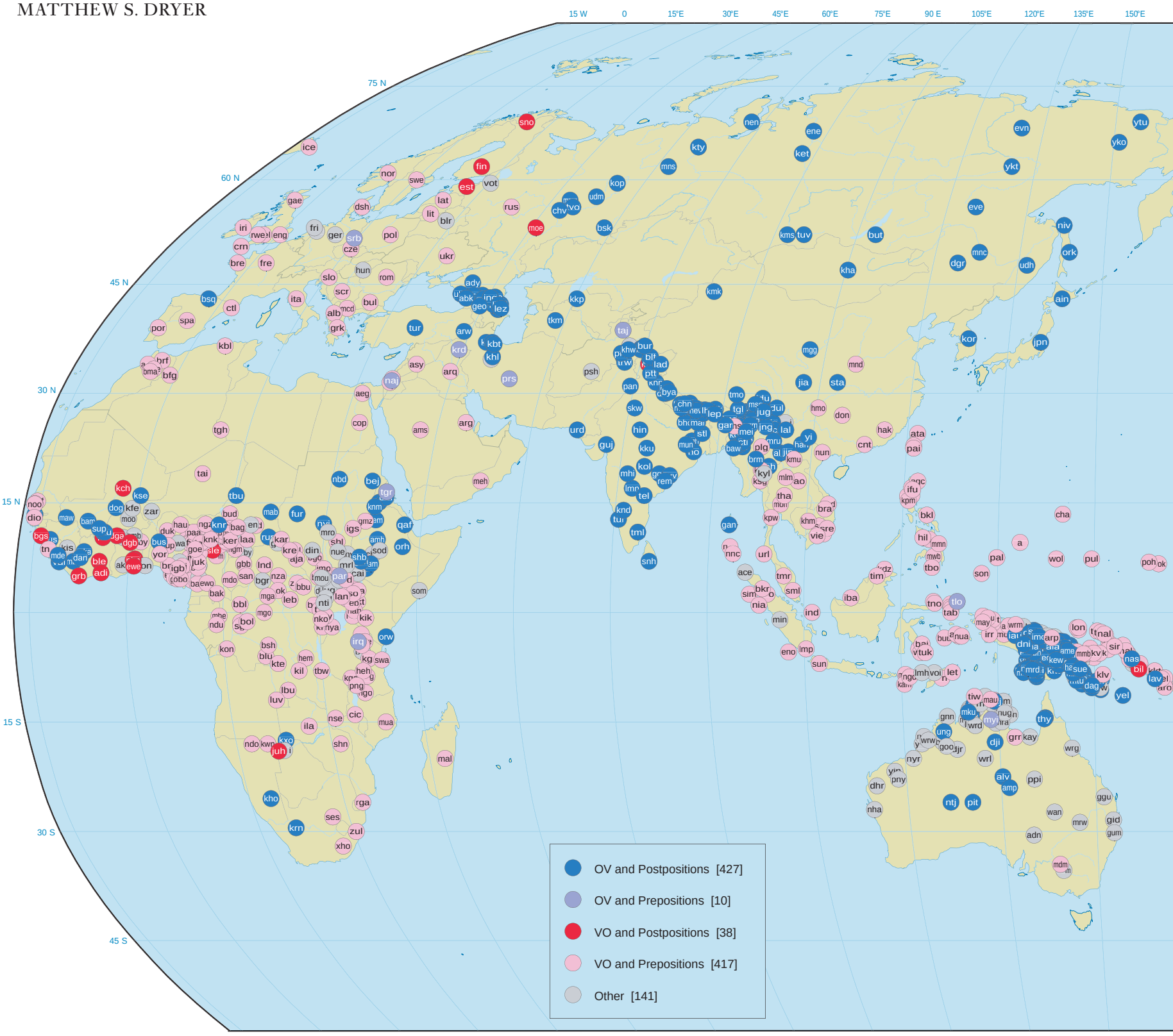
A number of the VO languages with postpositions are also atypical in that they have prepositions, too, although the postpositions are deemed the dominant type in the language. For example, Koyra Chiini, illustrated above in (3) as a VO language with postpositions, actually has a few prepositions in addition to its more numerous postpositions. The example in (9) illustrates the preposition *hal* 'until'.

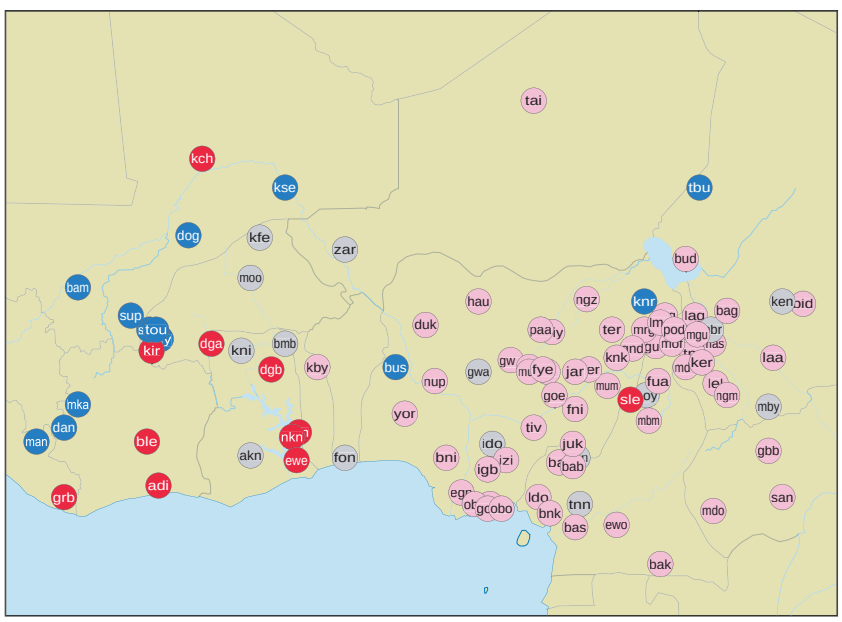
(9) Koyra Chiini (Heath 1999b: 108)
yer o nan ga hal alaasar
1PL IMPF leave 3SG.OBJ until afternoon
'We will leave it until the afternoon (prayer).'

In a number of cases, there is evidence that languages of one of the uncommon types have undergone a recent historical change in the order of object and verb and that the adposition type reflects the earlier word order. For example, Tigre, illustrated as an OV prepositional language in (2b) above, is one of a number of South Semitic languages in Ethiopia whose original order was clearly VO, based on comparative evidence, but which have undergone considerable structural change due to contact with Cushitic languages, which are OV. South Semitic languages vary as to which of their earlier word-order characteristics they have retained and which they have lost; in the case of Tigre, it has retained prepositions, but has undergone a change from VO to OV. It appears to be the case that if a language changes the order of one of these two features, it will almost always be the order of object and verb that changes first.

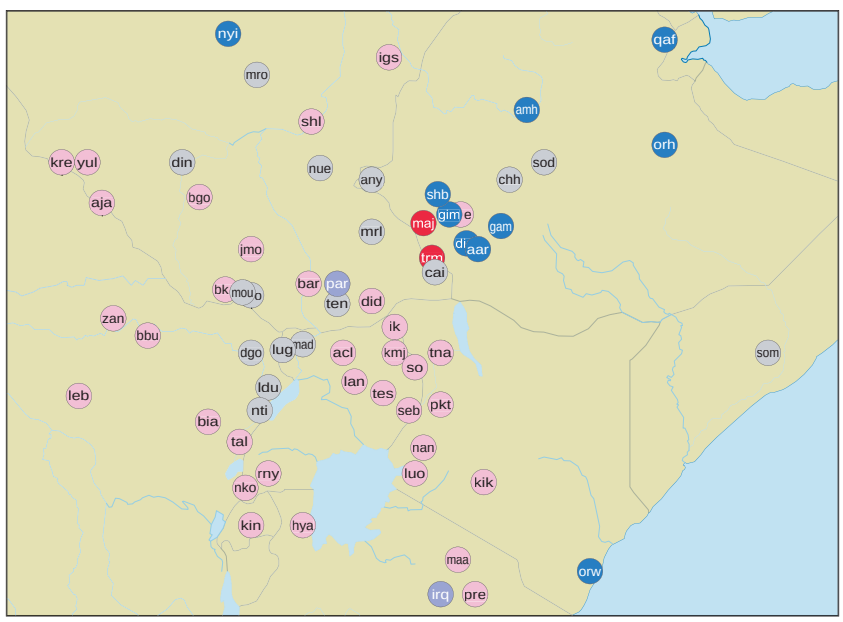
95 Relationship between the Order of Object and Verb and the Order

MATTHEW S. DRYER



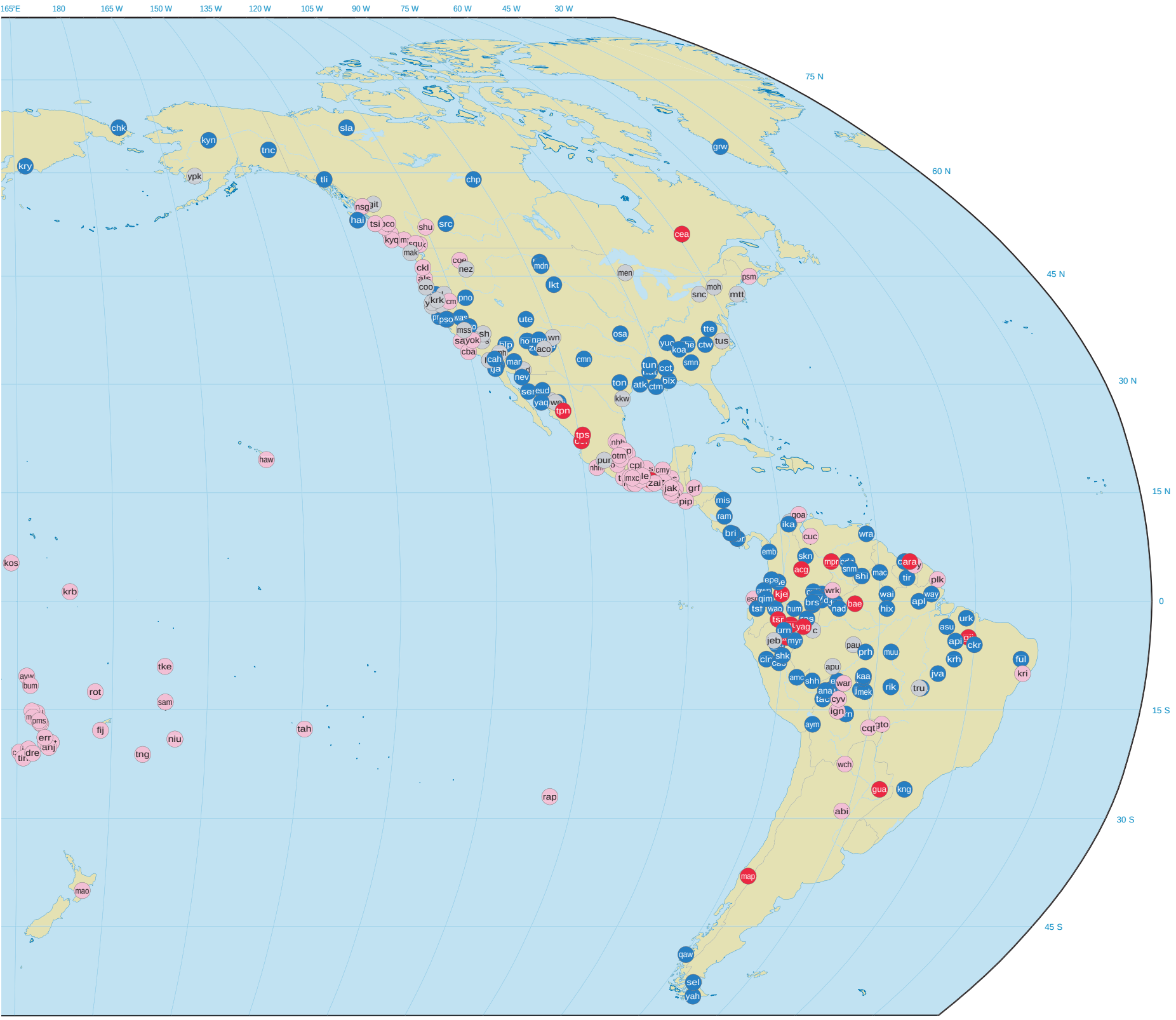


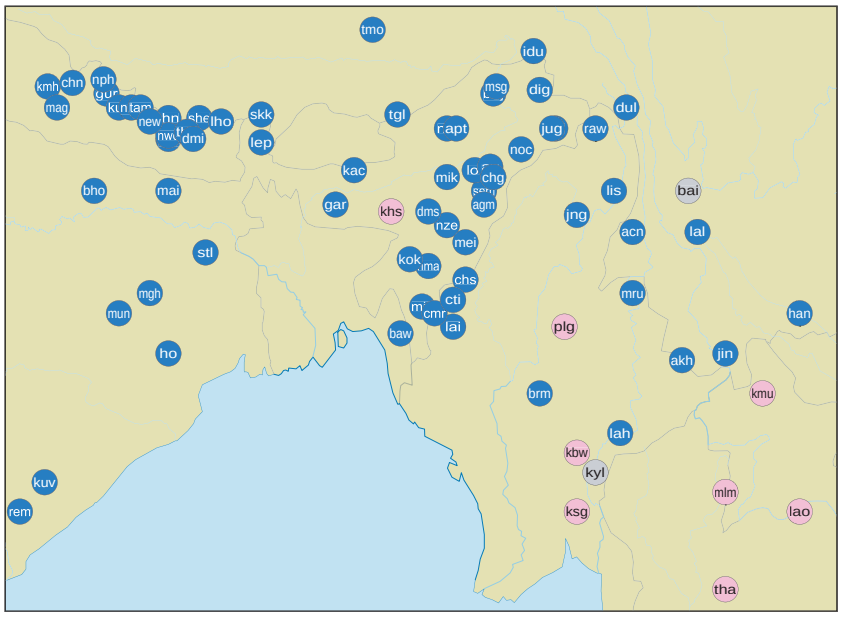
Nigeria and surrounding areas



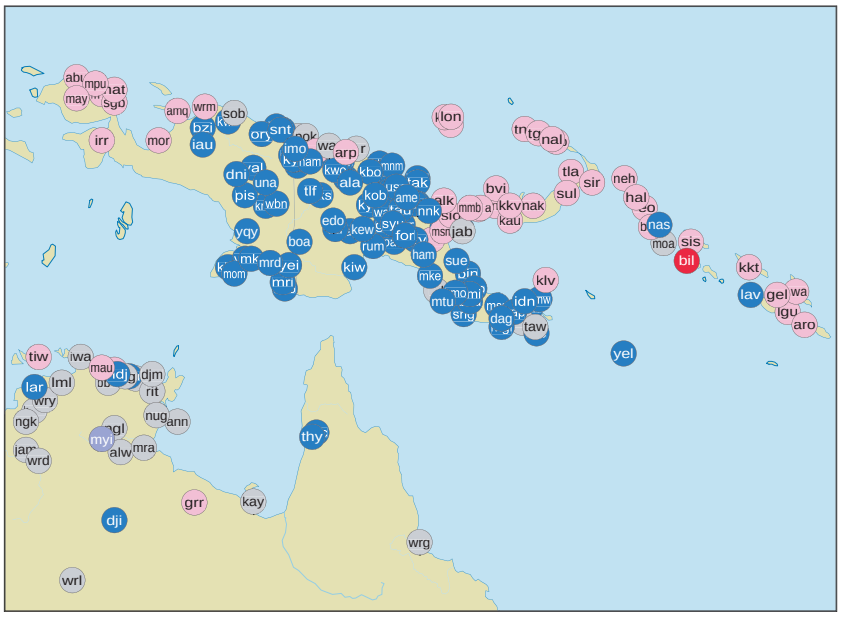
eastern Africa

# of Adposition and Noun Phrase





from Nepal to Thailand



New Guinea and northern Australia

96 Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun

MATTHEW S. DRYER

1 Defining the values

This is the second of three maps that show the relationship between two features that are shown separately in earlier maps, namely the order of relative clause and noun (Map 90) and the order of object and verb (Map 83). The two orders of object and verb and the two orders of relative clause and noun intersect to define the first four values shown in the box.

1.	Object–verb and relative clause–noun (OV&RelN)	109
2.	Object–verb and noun–relative clause (OV&NRel)	96
3.	Verb–object and relative clause–noun (VO&RelN)	5
4.	Verb–object and noun–relative clause (VO&NRel)	370
5.	Languages not falling into one of the preceding four types	176
	total	756

This map illustrates a case where three of the four types are common and one type is uncommon: the first, second, and fourth types are common, while the third is uncommon.

The first type consists of **OV and RelN** languages, languages in which the object precedes the verb and the relative clause precedes the head noun. An example of such a language is Kolyma Yukaghir (isolate; Siberia, Russia); (1a) illustrates the OV order while (1b) illustrates RelN order.

(1) Kolyma Yukaghir (Maslova 2003: 89, 146)

a. *met* *es'ie* *tet* *pulut-kele* *kudede-m*
1SG father 2SG husband-ACC kill-TR.3SG
O V
'My father has killed your husband.'

b. *[mit* *šohu-še-l]* *ani-pe*
[1PL get.lost-CAUS-ATTR.1PL] fish-PL
Rel N
'the fish [that we have lost]'

Similarly, the examples in (2) illustrate OV and RelN orders respectively in Rumu (Turama-Kikorian, Trans-New Guinea; Papua New Guinea).

(2) Rumu (Petterson 1999: 24, 19)

a. *ëkë* *po* *i* *te* *wë* *ru-nanë*
starling TOP tree LOC nest build-and.then
'The starlings build their nests in a tree.'

b. *[rö* *te* *ra-kë]* *uki* *reiapai*
[stern LOC stand-up] man one
'one man [who is standing up in the stern]'

Languages of the first type are common, though as the map shows and as is discussed below, they are much more common in Asia than in other parts of the world.

The second type consists of **OV and NRel** languages. While such languages are sometimes thought to be uncommon, it is clear from the map that they are in fact very common. An example of such a language is Harar Oromo (Cushitic; Ethiopia); the OV order is illustrated in (3a), the NRel order in (3b).

(3) Harar Oromo (Owens 1985: 107, 131)

a. *an* *gaalá* *xiyyá-n* *arke*
1SG camel my-1SG saw
'I saw my camel.'

b. *dubar-tíi* *[xennáa* *xann-ée-f]*
girl-NOM [present give-PST-DAT]
'the girl [he gave a present to]'

A second example of an OV and NRel language is Yaqui (Uto-Aztecan; northern Mexico); (4a) illustrates the OV order, (4b) a relative clause following the noun.

(4) Yaqui (Dedrick and Casad 1999: 40, 383)

a. *hunáa'a* *yá'uraa* *hunáka'a* *hámut-ta*
that chief that.ACC woman-ACC
nok-ria-k
speak-APPL-PERF
'Those authorities defended that woman.'

b. *húu'u* *'óo'ou* *[']ém* *bít-bae-'u]*
that man [2.POSS see-DESID-GERUND]
'the man [that you want to see]'

The third type consists of languages which are **VO and RelN**. An example of such a language is Mandarin, illustrated in (5), (5a) showing the VO order, (5b) the RelN order.

(5) Mandarin (Li and Thompson 1981: 128, 117)

a. *tāmen* *tōu* *zìxíngchē*
3PL steal bicycle
'They steal bicycles.'

b. *[wǒ* *gěi* *nǐ* *de]* *shū*
[1SG give 2SG LINK] book
'the book [that I gave you]'

Languages of this third type are distinctly rare. Apart from Mandarin and other varieties of Chinese, the only other cases on the map are two languages in close geographical proximity to Chinese, namely Bai, a Tibeto-Burman language of China that has been heavily influenced by Chinese, and Amis, an Austronesian language of Taiwan (Joy Wu, personal communication), where there is also a possibility of influence from Chinese. The examples in (6) show this for Bai, (6a) illustrating VO word order, (6b) the RelN order.

(6) Bai (Xu and Zhao 1984: 77, 73)

a. *a*^{31}*ti*^{33} *tshi*^{55} *tchi*^{55}
grandpa add fertilizer
'Grandpa add(ed/s) fertilizer (to the field).'

b. *[vẽ*^{42} *tse*^{21}*tsa*^{42} *no*^{33}*]* *s*^{55}
[write tidy LINK] word
'words [that are written tidily]'

The fourth type consists of languages which are **VO and NRel**. Examples of languages of this type include English, as in (7), and Hatam (West Papuan; Indonesia), as in (8).

(7) *the book [that I bought]*

(8) Hatam (Reesink 1999: 78, 115)

a. *i-ngot* *igya*
3PL-tie house
'They are building a house.'

b. *nab* *[po* *di-cig* *pek* *mem* *da]*
pig [REL 1SG-father buy for 1SG]
'the pig [that my father bought for me]'

Most VO languages are of this type.

The fifth type shown on the map includes various types of languages not falling into one of the first four types. This includes

languages lacking a dominant order of object and verb; languages with one of the four less frequent types of relative clauses shown on Map 90, namely internally headed relative clauses, correlative clauses, adjoined relative clauses, and double-headed relative clauses; and languages which have two or more of the relative clause types without one being dominant. Although not shown on the map, it is worth mentioning that these four less frequent types of relative clause are mostly found in OV languages. In fact, my data contains only one exception to this generalization, Kutenai (isolate; western North America), illustrated in (9), which is VO, in fact verb-initial, as in (9a), with internally headed relative clauses, as in (9b).

(9) Kutenai (own data)

a. *qa·ɬ* *ʍaqayq-ni* *ɬaǰanxuʔnaɬ-s* *niʔ*
in.that.way roll.up-IND door-OBV the
tiɬnamu
old.woman
'The old woman rolled up the teepee door.'

b. *nˀ=ip-s-i* *[niʔ-s* *k=ʍu·kat*
IND=die-OBV.SUBJ-IND [the-OBV SUBORD=see
paɬkiy-s *misaɬ]*
woman-OBV Mike]
'The woman that Mike saw died.'

One other VO language with internally headed relative clauses is Tukang Besi (Austronesian; Sulawesi, Indonesia), but this is not the dominant type of relative clause in the language, since it co-exists with externally headed postnominal relatives (Donohue 1999: 304, 367).

2 Geographical distribution

OV and RelN languages are very common in much of Asia except in an area extending from the Middle East to around Afghanistan and in eastern China and South-East Asia. Outside this area, they are not common, although there are multiple instances in (i) Ethiopia and Eritrea and (ii) New Guinea; in both of these areas, however, there are also many languages that are OV and NRel. OV and NRel languages are common in parts of North America, in northwestern South America, in those areas in Africa where OV languages are common, in an area in Asia stretching from Iraq to Afghanistan, and in Australia. In effect, the distribution of the two orders of relative clause and noun among OV languages is largely determined geographically: OV languages in Asia are generally RelN, while OV languages elsewhere in the world are much more commonly NRel, though there are small areas, as noted, in which RelN languages are common. It is often assumed, because of the high frequency of RelN order among OV languages in Asia, that it is normal for OV languages to be RelN; however, the evidence here shows that much of the rest of the world differs from Asia in this respect.

Since most VO languages are NRel, the distribution of VO and NRel languages on the map is similar to the distribution of VO languages on Map 83: they are common in (i) Europe and around the Mediterranean; (ii) much of sub-Saharan Africa; (iii) an area stretching from South-East Asia eastward through the Pacific; and (iv) Mesoamerica. Languages which are VO and RelN are confined to mainland China and Taiwan. Note that the location of these VO and RelN languages is adjacent to the area in which OV and RelN is most common, in Asia, and it seems likely that the occurrence of such languages in Asia is related to the overall high frequency of RelN order in Asia.

This map resembles in some respects the distribution on the next map, Map 97, which shows the relationship between the order of object and verb and the order of adjective and noun. The distribution of the two OV types on this map resembles the distribution of the two OV types on that map. On both maps, the type combining OV with AdjN or RelN is found in a large area covering most of Asia, and in roughly coterminous pockets elsewhere in the world. Apart from the fact that there are fewer languages shown on this map than on Map 97 (because grammatical descriptions more often give information on the order of adjective and noun), the most noticeable differences are (i) the fact that languages in the area around Myanmar are part of the large area in Asia in which OV and RelN order is common, but not part of the area in which OV and AdjN is common; and (ii) while there are many OV and AdjN languages in the southwestern United States, this is not a region in which OV and RelN is common. On the other hand, while the distribution of the two OV types is similar on the two maps, the distribution of the two VO types is quite different: on this map, almost all the VO languages are NRel, while on Map 97 both AdjN and NAdj are common among VO languages.

3 Theoretical discussion

The distribution on this map can be described in terms of an implicational statement "If a language is VO, then it is usually NRel", or by what is an essentially equivalent statement "If a language is RelN, then it is usually OV". Note that this implication is **unidirectional**, in contrast to that discussed in Chapter 95, where the implication is bidirectional. In other words, one cannot reverse the above implicational statements; one cannot say "If a language is NRel, then it is usually VO" or "If a language is OV, then it is usually RelN", both of which are false. It is true that among NRel languages, there are more that are VO, by 370 languages to 96, but this preference is weak compared to the huge preference for VO languages to be NRel, by 370 languages to 5.

The fact that one can formulate implicational generalizations relating these two typological features means that there is an interaction or correlation between them. This correlation is one of many correlations between the order of object and verb and the order of other pairs of elements, and the attempts at explaining these correlations, discussed in Chapter 95, are relevant to the correlation discussed in this chapter as well, although it is not clear that there is any good account in terms of grammaticalization for this correlation. In addition, since relative clauses are not complements of the noun, accounts in terms of consistent ordering of heads and complements will not account for the correlation found with relative clauses. Accounts in terms of heads and dependents will work, but will not account for the existence of a correlation with relative clauses but not with adjectives (see Chapter 97).

Any explanation for the correlation must also account for the unidirectionality of the implication. One type of approach is in terms of competing motivations; on this approach, there are explanatory forces that favour the type VO and NRel and the type OV and RelN, and other explanatory forces favouring NRel over RelN and thus favouring the type VO and NRel and the type OV and NRel. These forces would work together to favour NRel order among VO languages, but would be in competition with each other in OV languages, where one would favour OV and RelN, while the other would favour OV and NRel. Proposals by Hawkins (1990, 1993, 1994, 2002) are accounts of this sort.

The absence of a correlation between the order of object and verb and the order of adjective and noun, while there does exist a correlation between the order of object and verb and the order of relative clause and noun, might seem surprising in light of the fact that relative clauses resemble adjectives in terms of their relationship to the nouns they modify, as reflected by the fact that relative clauses are often referred to as **adjective clauses**. It is proposed in Dryer (1992) that the fact that we get a correlation only with relative clauses is due to the fact that they are phrasal, and in general contain more than one word, while adjectives are generally nonphrasal, consisting of a single word. Placing relative clauses before the noun in an OV language will tend to make the language consistently left-branching, while the order of adjective and noun will not affect the direction of branching in a language.

96 Relationship between the Order of Object and Verb and the Order

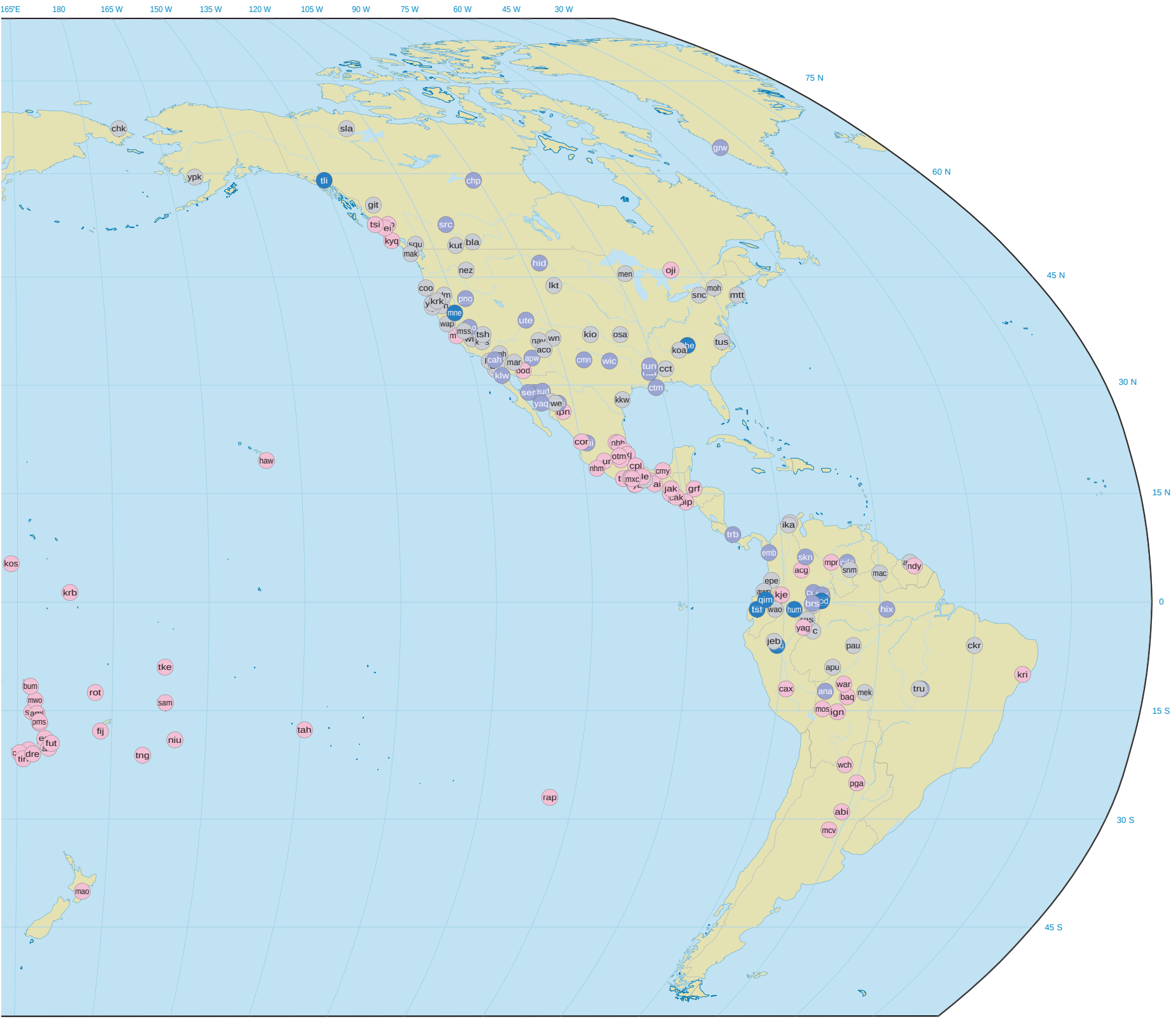
MATTHEW S. DRYER

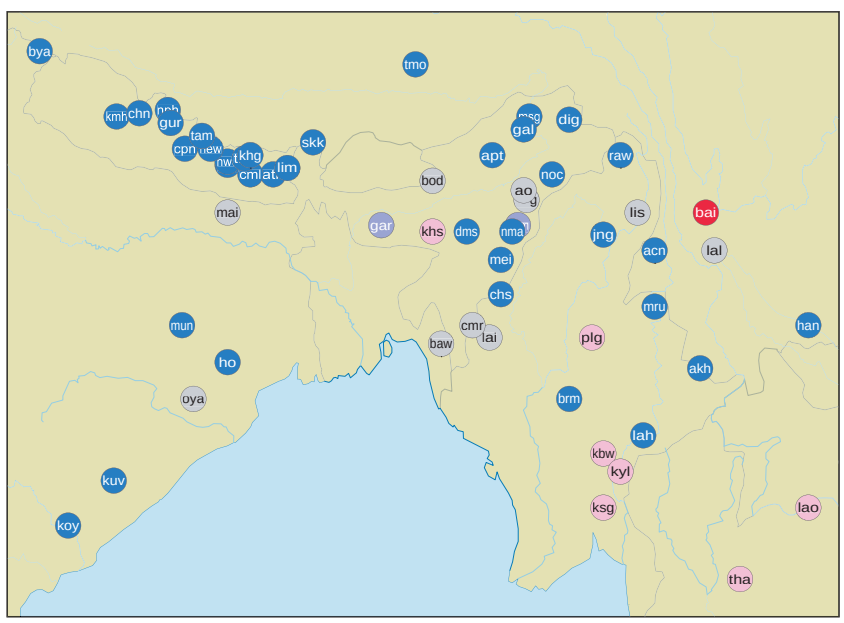
OV and RelN [109]
OV and NRel [96]
VO and RelN [5]
VO and NRel [370]
Other [176]

eastern Africa

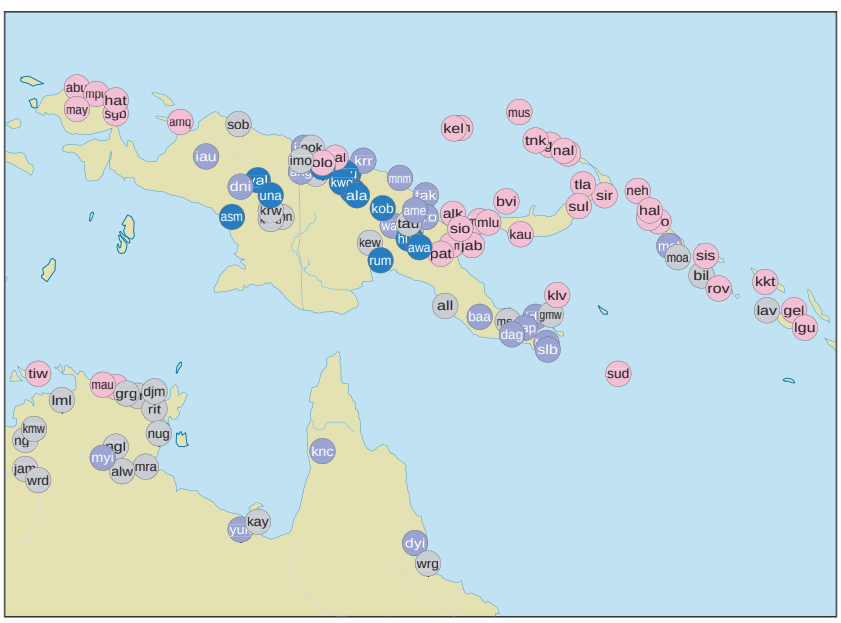
Caucasus and Middle East

# of Relative Clause and Noun





from Nepal to Thailand



New Guinea and northern Australia

97 Relationship between the Order of Object and Verb and the Order of Adjective and Noun

MATTHEW S. DRYER

1 Defining the values

This map shows the relationship between the order of object and verb and the order of adjective and noun; these two features are shown on Maps 83 and 87. This map is the third map showing the relationship between two typological features. Unlike the two previous maps, however, there is no correlation between the features or implicational relationship relating them: all four types defined on the intersection of these two features, the first four types shown in the box, are common.

1. Object–verb and adjective–noun (OV&AdjN)	201
2. Object–verb and noun–adjective (OV&NAdj)	287
3. Verb–object and adjective–noun (VO&AdjN)	100
4. Verb–object and noun–adjective (VO&NAdj)	404
5. Languages not falling into one of the preceding four types	178
total	1170

The first type, **OV and AdjN**, languages in which the object precedes the verb and the adjective precedes the noun, is illustrated in (1) from Udihe (Tungusic; Siberia), (1a) illustrating the OV word order, (1b) the AdjN order.

(1) Udihe (Nikolaeva and Tolskaya 2001: 840, 487)

a. *mamasa* *ule:-we* *olokto-ini*
old.woman meat-ACC cook-3SG
O V
'The old woman is cooking meat.'

b. *uligdig'a-ŋku* *moxo-ziga*
beautiful-PL cup-PL
Adj N
'beautiful cups'

The examples in (2) illustrate a language which is **OV and NAdj**, Kunuz Nubian (Nilo-Saharan; southern Egypt).

(2) Kunuz Nubian (Abdel-Hafiz 1988: 219, 283)

a. *id* *ti:-g* *kanni:r-ken* *gos-s-u*
man cow-ACC knife-INSTR slay-PST-3SG
'The man slaughtered the cow with the knife.'

b. *e:n* *duru-we:r*
woman old-INDEF
'an old woman'

The third type is languages which are **VO and AdjN**. This type is illustrated by English, as well as by Coast Tsimshian (Tsimshianic; British Columbia); the VO order is illustrated in (3a), the AdjN order in (3b).

(3) Coast Tsimshian (Mulder 1994: 32, 162)

a *yagwa-t* *huum=[da* *duus]=[a* *hoon]*
PRES-3.ERG smell=[ERG cat]=[ABS fish]
'The cat is sniffing the fish.'

b. *'wii* *samn*
great spruce.tree
'a big spruce tree'

(Note that the case markers in Coast Tsimshian are prenominal but encliticize onto the preceding word; for example the absolutive case marker =*a*, which marks the object noun phrase *hoon* 'fish' as absolutive, attaches phonologically to *duus* 'cat', but grammatically it goes with *hoon* 'fish'; the square brackets indicate the syntactic constituents.)

The fourth type is languages which are **VO and NAdj**. This type is represented by Olo (Torricelli; Papua New Guinea); both VO and NAdj orders are illustrated in (4).

(4) Olo (Staley 1995: 31)

ki *k-alei* *tifa* *funi*
1SG 1SG.SUBJ-eat.3SG M.OBJ banana ripe
'I eat a ripe banana.'

The fifth type shown on the map involves **languages not falling into one of the first four types**. It includes languages lacking a dominant order of object and verb or a dominant order of adjective and noun, as well as languages lacking a construction in which adjectives modify nouns because the closest equivalent involves internally headed relative clauses (see Chapter 87).

2 Geographical distribution

Languages which are OV and AdjN are especially common in much of Asia, except in the Middle East and in an area stretching from northeastern India through South-East Asia. Elsewhere in the world, this type is found scattered and in various small areas, though in many of these it is interspersed with OV and NAdj languages. These areas include (i) Ethiopia and Eritrea; (ii) two areas in Papua New Guinea, one in the eastern highlands, the other in the lower Sepik valley; (iii) an area in the southwestern United States and northwestern Mexico; and (iv) an area in Ecuador and adjacent parts of Brazil to the east.

Languages which are OV and NAdj are the dominant type in Australia and in New Guinea. They are also the dominant type in North America outside two well-defined areas in the Pacific North-West and Mesoamerica; of the four main types, they are the primary type represented in Canada and the United States east of the Rocky Mountains as well as in the north, including Alaska, northern Canada, and Greenland. They are also the most frequent type in South America, though other types are common and the four types do not show clear areal patterning within South America. They are also found in (i) Myanmar and adjacent areas in India and China (all the languages in question being Tibeto-Burman); and (ii) various parts of Africa where OV languages are spoken.

Languages which are VO and AdjN are dominant in (i) an area in northern and eastern Europe; and (ii) the Pacific North-West. There are also smaller pockets in (i) an area in central Africa centred in the Central African Republic; (ii) Taiwan and the Philippines; (iii) the extreme northern part of Northern Territory in Australia; (iv) an area in southeastern Australia; (v) two pockets in Mesoamerica, one in the northern part of this region, the other in the south-east. This type is also represented by the different varieties of Chinese, though these do not show up on the map as a well-defined area.

Languages which are VO and NAdj are the dominant type in (i) western Europe and around the Mediterranean, including the Arabian Peninsula; (ii) sub-Saharan Africa; (iii) a large area stretching from South-East Asia eastward through the Pacific, except in the Philippines and New Guinea. They are also common in Mesoamerica, though there are also many VO and AdjN languages in this area, especially to the north and south-east.

3 Theoretical issues

The order of adjective and noun is often claimed to correlate with the order of object and verb. Dryer (1988a, 1992) argues against

this, however, claiming that AdjN order occurs as often in VO languages as it does in OV languages. It has been known since Greenberg (1963) that many OV languages are NAdj. What Dryer (1988a, 1992) argues, however, is (i) that NAdj order is even more common among OV languages than had been realized, AdjN order being more common only in Eurasia; and (ii) that AdjN order is more common in VO languages than previously thought.

The first of these, that AdjN is more common among OV languages only in Eurasia, is reflected by the map. Nevertheless, the map somewhat exaggerates the frequency of OV and AdjN languages, for two reasons. First, the projection used shows areas further from the equator as somewhat larger than they really are. As a result, the area across northern Asia appears larger relative to equatorial regions than it actually is. Second, there are more languages and greater genealogical diversity in regions closer to the equator, so that there are a greater number of languages and genealogical groups (families and genera) with OV and NAdj languages than is reflected on the map. There are more languages in New Guinea, for example, than there are on the mainland of Europe and Asia combined. Similarly, there are more than twice as many languages in Africa as there are on the mainland of Europe and Asia combined.

In terms of raw numbers of languages, AdjN order is less common among VO languages than it is among OV languages. Similarly, in terms of raw numbers of languages, languages which are VO and NAdj are considerably more common on the map than languages which are VO and AdjN. But this is misleading. It reflects the fact that there are two very large language families, Niger-Congo and Austronesian, each containing about 20 per cent of the languages of the world, in which the majority of languages are VO and NAdj. More than half of the languages shown on the map which are VO and NAdj (221 out of 403) are in one of these two families: the vast majority of the VO and NAdj languages in sub-Saharan Africa are Niger-Congo, and most of the languages of this type in the region stretching from Indonesia eastward into the Pacific are Austronesian. If one counts, not number of languages, but number of genera—genealogical groups comparable to the subfamilies of Indo-European (see the Introduction to the Genealogical Language List)—containing languages of the various types, then the number of genera containing languages which are VO and NAdj is indeed greater than the number that are VO and AdjN. However, the ratio of genera containing NAdj languages to those containing AdjN languages is only slightly greater for VO languages than it is for OV languages. The following table gives the number of genera containing languages of the type given:

(5)

OV&AdjN	70
OV&NAdj	136
VO&AdjN	51
VO&NAdj	95

The ratio of VO&NAdj to VO&AdjN is 1.86, which is actually less than the ratio of OV&NAdj to OV&AdjN, which is 1.94. Furthermore, if one counts the number of language families containing languages of each of the four types, assuming the highest level of classification in the Genealogical Language List as a criterion for what constitutes a family, we get the numbers in the following table:

(6)

OV&AdjN	43
OV&NAdj	61
VO&AdjN	28
VO&NAdj	33

Here we find considerably more families containing OV&NAdj languages compared to OV&AdjN (61 families to 43), but only slightly more families containing VO & NAdj languages compared to VO&AdjN (33 families to 28).

The basis for the mistaken impression that many linguists have had that there is a correlation between the order of object and verb and the order of adjective and noun apparently derives from the fact that within Europe and Asia, there is the appearance of such a pattern. On the one hand, there is an overwhelming preference for AdjN order among the OV languages of Asia. Conversely, although there is a large area of VO languages in northern Europe with AdjN order, the VO languages of southwestern Europe are NAdj, as are a clear majority of the VO languages of South-East Asia. In addition, the fact that the majority of VO languages within two very large families, Niger-Congo and Austronesian, are NAdj adds to the impression.

In light of the above discussion, the data provide no basis for thinking that OV languages have any stronger preference for AdjN order than VO languages do. However, some have suggested (e.g. Greenberg 1963) that AdjN order is less frequent in verb-initial languages (as opposed to SVO languages, which like verb-initial languages are VO). But although this is not shown on this map, the languages used for the map do not support this hypothesis either. The numbers in (7) give the number of genera and number of families for the two orders of adjective and noun in verb-initial languages, analogous to the numbers given above in (5) and (6):

(7)

	No. of genera	No. of families
Verb-initial&AdjN	35	24
Verb-initial&NAdj	34	18

The data in (7) provides no reason to think that AdjN order is any less frequent in verb-initial languages. If anything, AdjN order is more frequent in verb-initial languages, since AdjN outnumbers NAdj among verb-initial languages both in terms of numbers of genera (35 to 34) and in terms of numbers of families (24 to 18), while NAdj outnumbers AdjN both for OV languages and for VO languages in general (see (5) and (6) above). Examples of verb-initial languages which are AdjN include Coast Tsimshian, illustrated above in (3), Batad Ifugao (Austronesian; Philippines), illustrated in (8), and Gude (Chadic; Nigeria), illustrated in (9).

(8) Batad Ifugao (Newell 1993: 15, 23)

a. *in-dat* *Aligūyun* *nan* *dotag*
THEME.FOC:PST-give Aligūyun DEF meat
ay *agī-na*
LOC brother-3SG
'Aligūyun gave the meat to his brother.'

b. *nan* *nappuhi-n* *tibung*
DEF bad-LINK wine.jar
'the bad wine jar'

(9) Gude (Hoskison 1983: 90, 53)

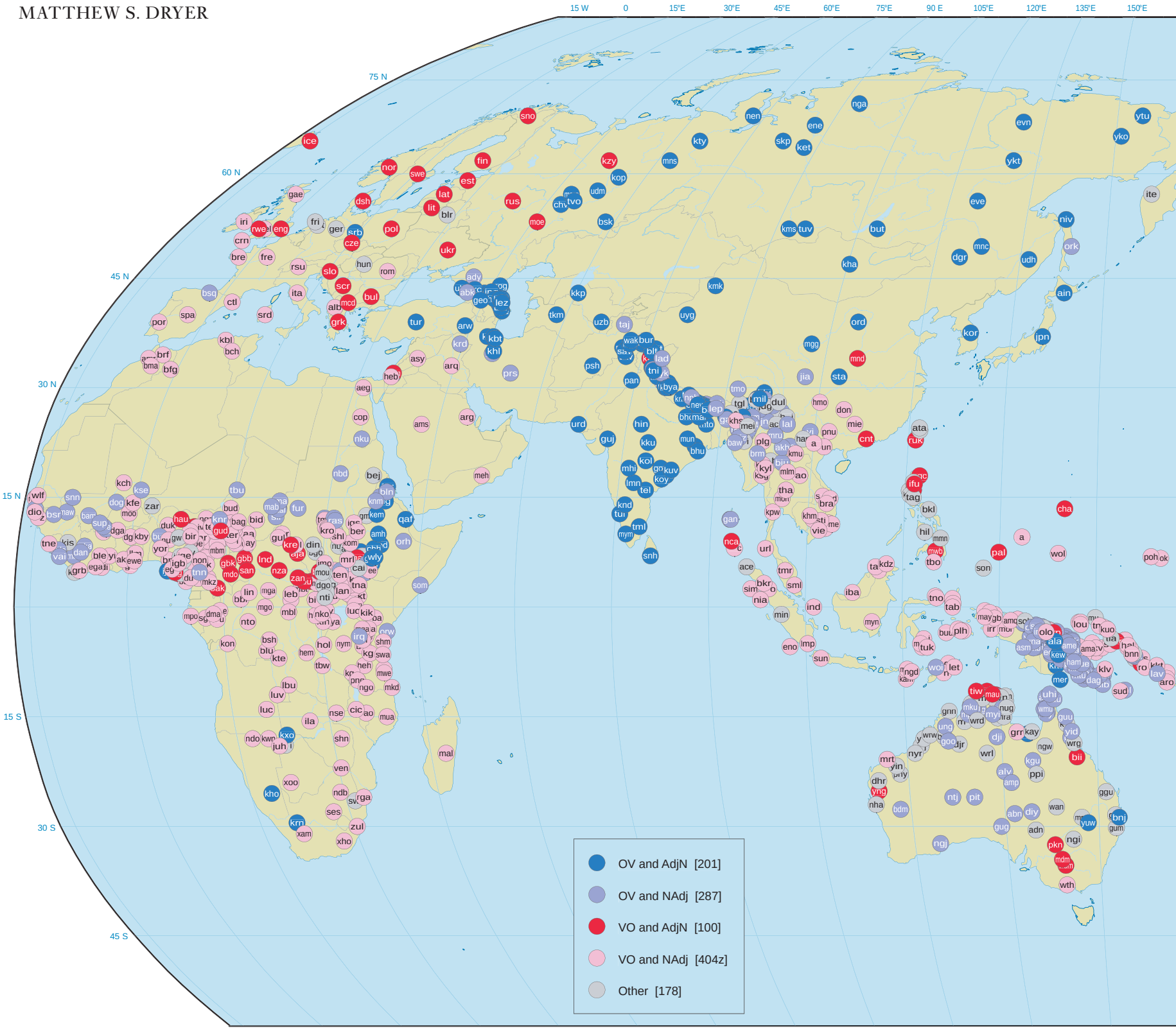
a. *kə* *kii* *Musa* *faara*
COMP throw Musa stone
'Musa threw a stone.'

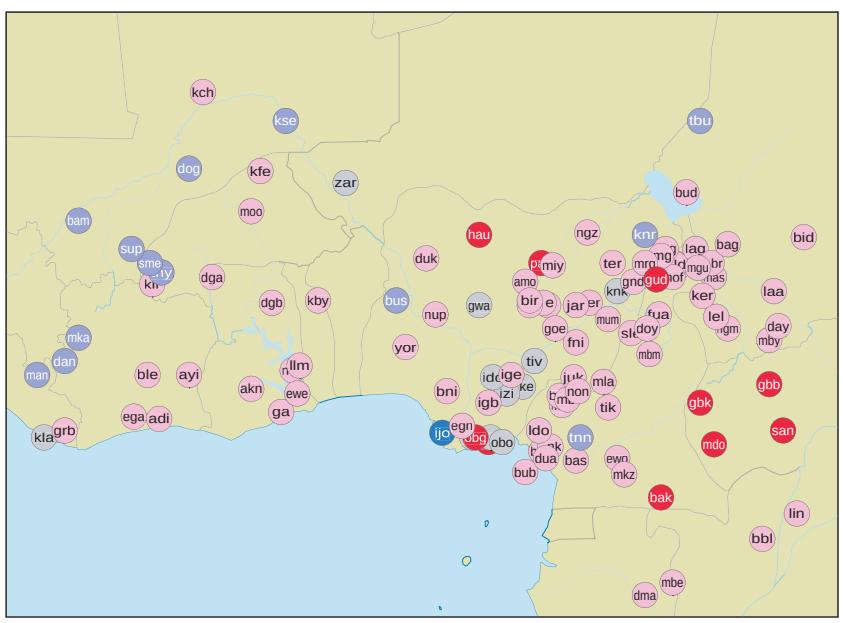
b. *ma-gusə* *minə*
ATTR-short woman
'a short woman'

As discussed in Chapter 87, the term "adjective" is used here in a semantic sense, for words denoting properties, regardless of their word class in different languages. In many languages, adjectives are formally verbs, and hence as modifiers of nouns they constitute a type of relative clause. This raises the question of whether distinguishing different languages on the basis of the word class of such adjectival words within the language might lead to new generalizations. Preliminary results suggest that this is not the case, but further investigation is necessary.

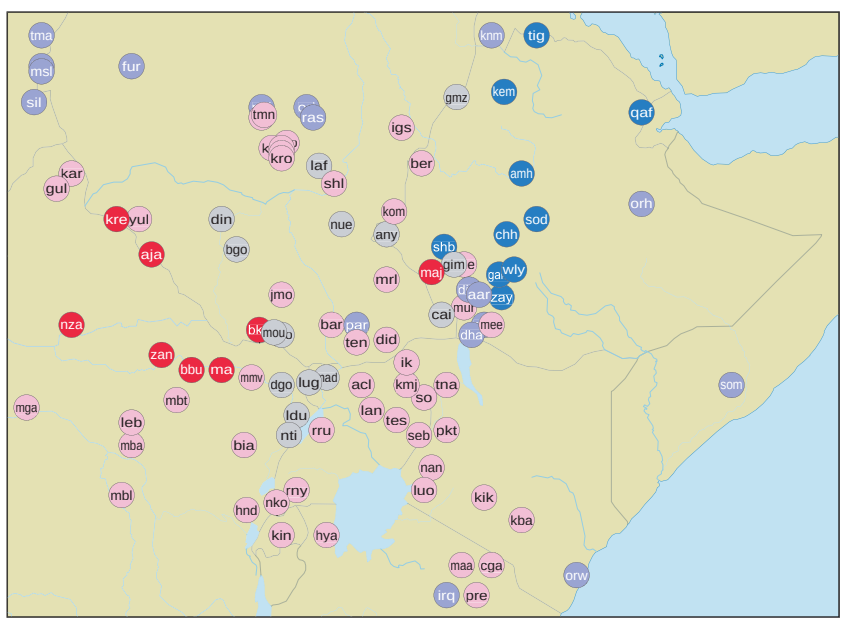
97 Relationship between the Order of Object and Verb and the Order

MATTHEW S. DRYER



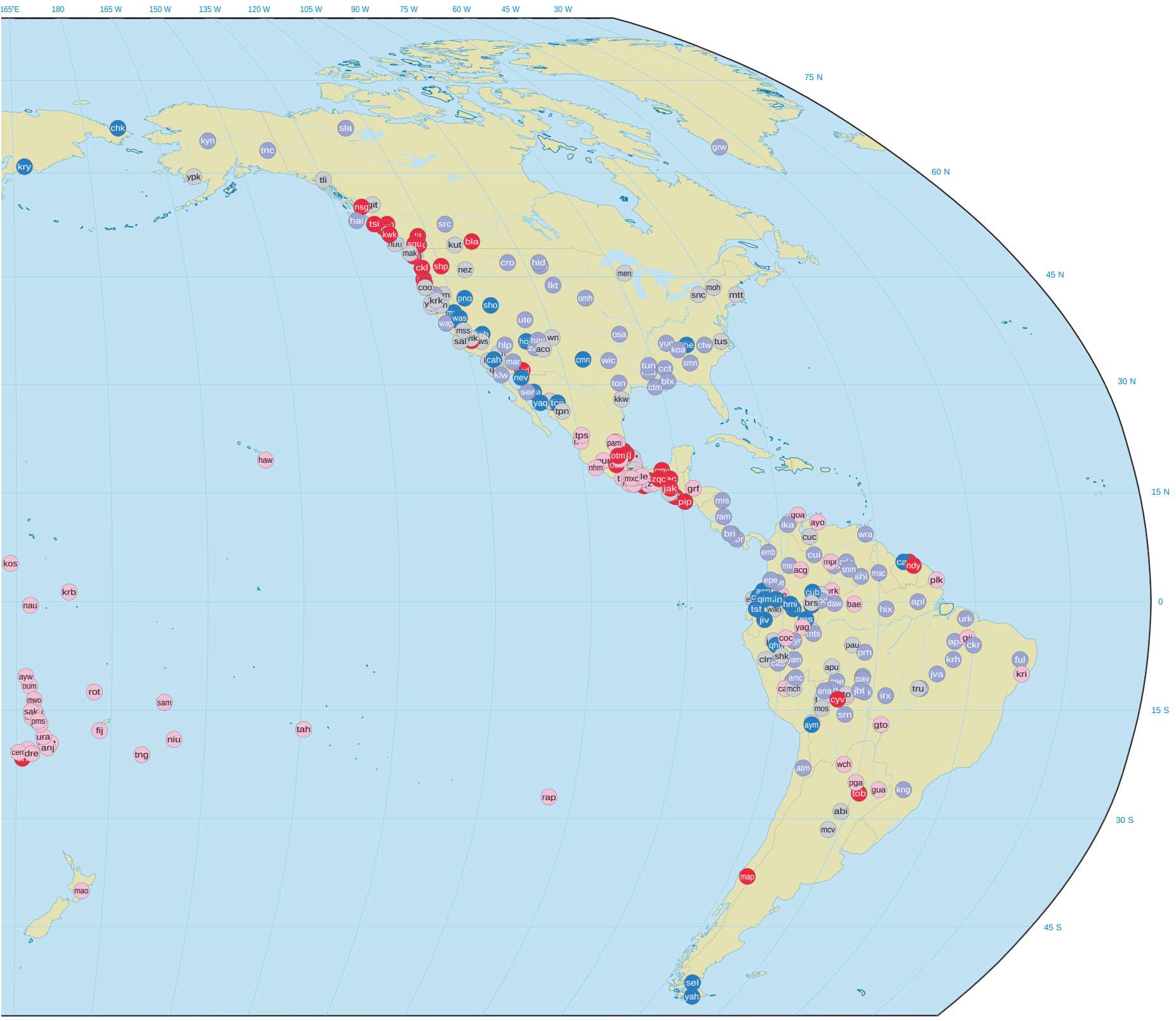


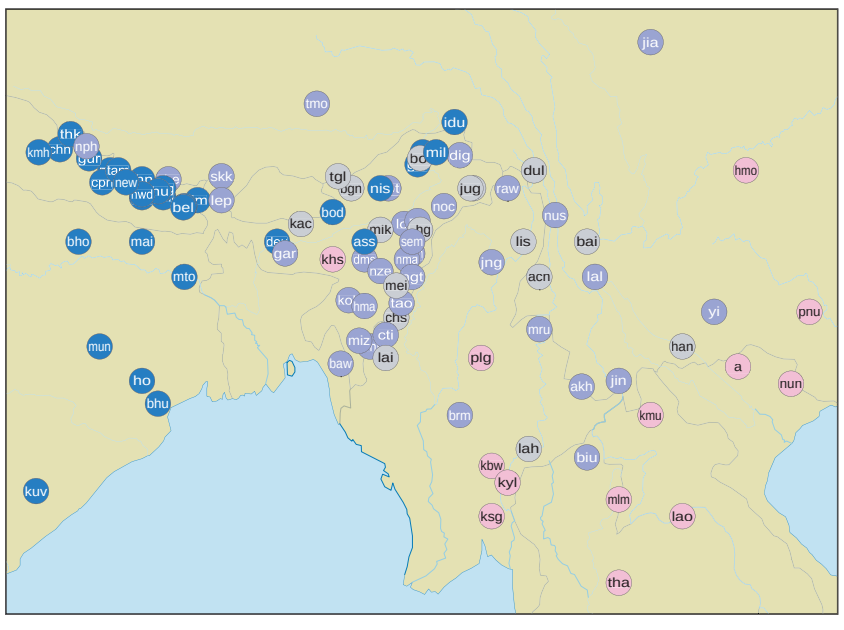
Nigeria and surrounding areas



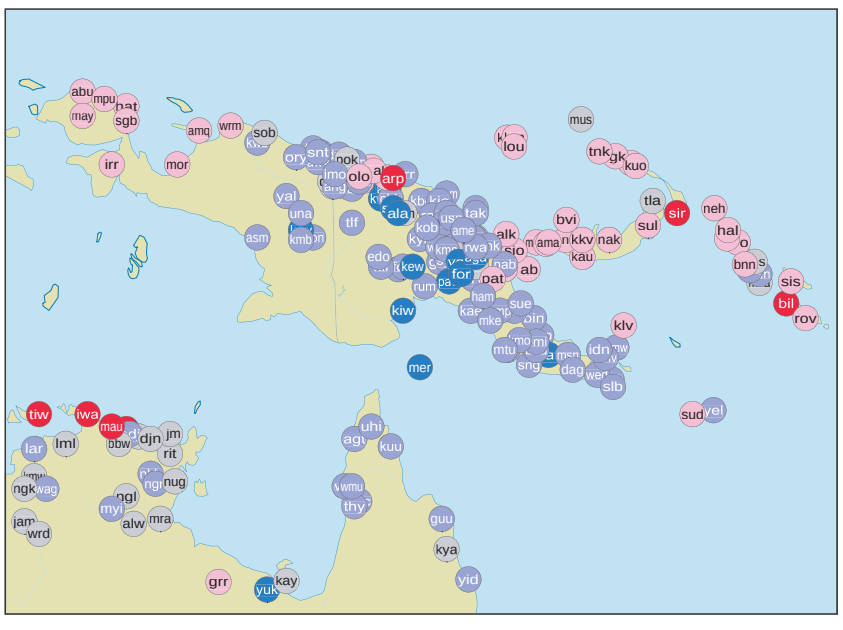
eastern Africa

# of Adjective and Noun





from Nepal to Thailand



New Guinea and northern Australia

98–99 Alignment of Case Marking

BERNARD COMRIE

1 Main alignment types

These two maps are concerned with the ways in which core argument noun phrases are marked—by means of morphological case or adpositions—to indicate which particular core argument position they occupy. The first map is concerned with full noun phrases, the second with pronouns.

The core argument of a canonical, one-place, intransitive predicate may be symbolized S. The two core arguments of a canonical, two-place, transitive predicate may be symbolized as A and P, with A representing the more agent-like argument and P the more patient-like (Comrie 1978). (In another terminology, the symbol O is used rather than P; Dixon 1994.) In studying the alignment of case marking, we ask the question which of S, A, and P are coded identically and which are coded differently. Note that for the purposes of this chapter, only case marking is considered. Alignment of person marking in the verb is treated in Chapter 100. Other manifestations of alignment are also possible, such as word order, but are not treated here.

In the **neutral** case-marking system, all of S, A, and P are marked in the same way. This can be illustrated by Mandarin examples (1a–b), where neither the S of (1a) ('the person'), nor the A of (2b) ('Zhangsan'), nor the P of (2b) ('Lisi') receives any case marking.

(1) Mandarin (Li and Thompson 1981: 20)

a. *rén* *lái* *le*
person come CRS
'The person has come.'

b. *zhāngsān* *mà* *lǐsì* *le* *ma*
Zhangsan scold Lisi CRS Q
'Did Zhangsan scold Lisi?'

In the **nominative–accusative** (or simply: accusative) case-marking system, S and A are marked in the same way, while P is marked differently. The form used to encode S and A is referred to as the *nominative*, the form used to encode P as the *accusative*, as illustrated in Latvian examples (2a–b).

(2) Latvian (Mathiassen 1997: 181, 187)

a. *Putn-s* *lidoja.*
bird-NOM fly.PST.3
'The bird was flying.'

b. *Bērn-s* *zīmē* *sun-i.*
child-NOM draw.PRES.3 dog-ACC
'The child is drawing a dog.'

Note that the definition of the nominative–accusative system says nothing about *how* the distinction between S/A and P is marked. In Latvian, both nominative and accusative have overt markers. However, it is also possible for just the accusative to have an overt marker, as in Hungarian, where the word for 'person' is *ember* in the nominative, but *ember-t* in the accusative. Much less frequently cross-linguistically, it is the nominative that has an overt marker and the accusative that lacks one, as in Harar Oromo (Cushitic, Afro-Asiatic; Ethiopia) examples (3a–b).

(3) Harar Oromo (Owens 1985: 101, 251)

a. *sárée-n* *adíi-n* *ní* *iyyi-t-i*
dog-NOM white-NOM FOC bark-F-IMPF
'The white dog is barking.'

b. *haat-tíi* *okkóttée* *goot-t-i*
mother-NOM pot make-F-IMPF
'Mother is cooking (lit. making the pot).'

Since the "**marked nominative**" type illustrated by Harar Oromo is a topic of current typological and theoretical interest, it has been given a separate encoding in the maps, contrasting with the standard type where either just the accusative or both nominative and accusative are marked.

In the **ergative-absolutive** (or simply: ergative) system, S and P are encoded in the same way, and A is encoded differently, as in Hunzib (Nakh-Daghestanian; eastern Caucasus); see examples (4a–b).

(4) Hunzib (van den Berg 1995: 122)

a. *kid* *y-ut'-ur*
girl CL2-sleep-PST
'The girl slept.'

b. *ožd-i-l* *kid* *hehe-r*
boy-ERG girl hit-PST
'The boy hit the girl.'

The case that encodes S and P is referred to as the *absolutive*, the case that encodes A as the *ergative*. (In an alternative terminology, the case that encodes S and P in the ergative–absolutive system is referred to as the *nominative*. This usage is not adopted here, to avoid confusion.) In Hunzib, the ergative case has an overt marker, *-l*, while the absolutive does not. However, it is also possible for both cases to have overt markers, as in Tukang Besi (Western Malayo-Polynesian; Indonesia; Donohue 1999: 51), where the preposition *na* marks the absolutive, the preposition *te* the ergative. The "marked absolutive" is exceedingly rare, having been so far attested only in one language, Nias (Western Malayo-Polynesian; Indonesia), where the absolutive is marked by modifying the initial segment of the ergative (Brown 2001).

In the **tripartite** system, all of S, A, and P are marked differently. This system is found for some noun phrases in Hindi, as illustrated in examples (5a–b).

(5) Hindi (Anvita Abbi, p.c.)

a. *laRkaa* *kal* *aay-aa*
boy yesterday come.AOR-SG.M
'The boy came yesterday.'

b. *laRke* *ne* *laRkii* *ko* *dekh-aa*
boy.OBL ERG girl ACC see-SG.M
'The boy saw the girl.'

In (5a), the S has no overt marker. In (5b), the A has the ergative postposition *ne* (which requires the preceding noun to be in the oblique case), while the P has the accusative postposition *ko*. For only one language has it been claimed that all noun phrases have the tripartite system, namely Warrungu (Pama-Nyungan; Australia; Tasaku Tsunoda, p. c.).

There is one other logical possibility for grouping S, A, and P, namely for A and P to have the same form, while S has a distinct form. This possibility is exceedingly rare; it does not occur in our sample, but is attested in some Iranian languages of the Pamir region, though restricted to some pronouns (Payne 1979).

In all of the systems discussed so far, there has been consistent encoding of all instances of S in the same way. However, another possibility is for S to be split between more agent-like and more patient-like instances of S, which we may symbolize as S_a and S_p respectively. On the basis of semantic similarity, S_a then groups with A, while S_p groups with P. This system has come to be called the **active–inactive** (or simply: active) system, on the basis of terminology originally created by the Russian linguist Georgij A. Klimov, though other terms are also found, e.g. *agentive–patientive* or *stative-active*. The active form covers S_a and A, the inactive S_p and P. This system is rather widespread as a basis for person marking on verbs (see Chapter 100), but is also found occasionally with case marking, as in examples (6a–c) from Georgian.

(6) Georgian (Harris 1981: 40)

a. *vaxt'ang-i* *ekim-i* *iqo*
Vakhtang-PAT doctor-PAT be.AOR.3SG
'Vakhtang was a doctor.'

b. *nino-m* *daamtknara*
Nino-AGT yawn.AOR.3SG
'Nino yawned.'

c. *nino-m* *ačvena* *surat-eb-i*
Nino-AGT show.AOR.3SG>3SG>3SG picture-PL-PAT
gia-s
Gia-DAT
'Nino showed the pictures to Gia.'

In (6a), the more patient-like S of the copular verb stands in the inactive case, while in (6b) the more agent-like S of 'yawn' stands in the active case. The active case is also used for the A of (6c), the inactive case for its P. The active–inactive system is rare for case marking, and is identified in only four languages of the sample for these chapters: Drehu (Oceanic; New Caledonia; Moyse-Faurie 1983: 147), Basque (Hualde and Ortiz de Urbina 2003: 364), Georgian, and Imonda (Border family; Papua New Guinea; Seiler 1985: 138–9). The available examples are consistent with the criteria distinguishing active from inactive clauses with respect to person marking on verbs, such as volitionality and dynamicity (Mithun 1991).

Finally, the system "**none**" is restricted to pronouns and is used for languages where pronouns are not permitted in one or more of the positions S, A, and P; in the sample used for the chapters these languages are Wari' (Chapacura-Wanham; Brazil; Everett and Kern 1997: 122) and Wichita (Caddoan; Oklahoma; Rood 1976: 10) (in both of which personal pronouns are not found for any of S, A, and P) and Canela-Krahô (Macro-Ge; Brazil; Popjes and Popjes 1986: 175), in which personal pronouns are found in S and A position, but not in P position.

2. Map 98: Alignment of case marking of full noun phrases

1.	Neutral	98
2.	Nominative–accusative (standard)	46
3.	Nominative–accusative (marked nominative)	6
4.	Ergative–absolutive	32
5.	Tripartite	4
6.	Active–inactive	4
	total	190

In deciding which type to assign a particular language to with regard to its case marking alignment, a number of problems arise, and as far as possible an attempt has been made to find consistent—or even better, principled—solutions to these problems, although a number of difficult cases remain. In this section, problems relating to full noun phrases are treated. Most of these same problems also carry over to pronouns, but pronouns introduce a further set of difficulties, which form the topic of §3.

First, there is a general problem concerning the dividing line between the ergative (including the tripartite) and active systems. In a number of languages that have a basically ergative system, a small number of intransitive verbs, or a small semantic range of intransitive verbs, require their S to be in the case identified as ergative. Such languages thus stand between a pure ergative system (where all intransitive verbs take their S in the absolutive) and a pure active system (where intransitive verbs are divided into two substantial subsets, one taking S in the active, the other taking S in the inactive). The policy adopted here is that for a language to be considered of the active type, there must indeed be two substantial sets of intransitive verbs differing in the case marking of their S. A small number of exceptional intransitive verbs, as in Hindi (McGregor 1977: 73–4), or a small semantic area of intransitive verbs (as with onomatopoeic verbs in Hunzib; van den Berg 1995: 124), will not be taken into account, so that Hindi is considered tripartite, Hunzib ergative. Sometimes, there are dialect differences within a language; thus, both Basque and Georgian dialects differ in how many intransitive verbs take active subjects.

But the main recurrent difficulty is that in many languages, different kinds of full noun phrases partake of different case-marking patterns. For instance, in Spanish the accusative marker, the preposition *a*, is found (roughly) only with specific, animate noun phrases, so that strictly speaking a noun phrase like the male proper name *Juan* has a nominative–accusative case-marking system, while the inanimate noun phase *el libro* 'the book' has a neutral case-marking system, as illustrated partially in (7).

(7) Spanish

a. *María* *vio* *a* *Juan*.
Mary see.AOR.3SG ACC John
'Mary saw John.'

b. *María* *vio* *el* *libro*.
Mary see.AOR.3SG the book
'Mary saw the book.'

(Instances where the P sometimes takes case marking, sometimes not, have come to be called *differential object marking*.) In some languages, a case marker is used primarily to avoid ambiguity, so that in Lower Grand Valley Dani (Trans-New Guinea; Papua, Indonesia; Bromley 1981: 84–5), for instance, the ergative marker is used in this way. In yet other languages, case markers may be described as optional, without any detailed discussion of the conditions under which the marker does or does not occur, e.g. the accusative marker in Burmese or the ergative marker in Araona (Tacanan; Bolivia; Pitman 1980: 14). More complex patterns may arise, e.g. in Gooniyandi (Bunaban; Australia) the ergative marker is optional with animate nouns but obligatory with inanimate nouns (McGregor 1990: 319–20); in Hindi, the accusative marker is used for Ps that are higher in definiteness and especially animacy, while the ergative marker is obligatory for A. The policy that has been followed in assigning such languages to types has been to maximize the occurrence of overt case marking. Thus, if a language has an optional accusative case marker, or one that occurs only under certain specified circumstances, then this has been given priority and taken as criterial. This policy decision needs to be taken into account consistently in interpreting the maps. (For details on how decisions were taken for individual languages, reference should be made to the electronic version of this atlas.) Thus, Spanish and Burmese come out as accusative, Araona and Gooniyandi as ergative, and Hindi as tripartite.

In a number of languages, the case-marking system is different in different tense-aspect-moods (TAMs). In Georgian, there are three sets of TAMs with respect to the fine details of case marking, grouping into two sets with respect to the factors relevant to present concerns. In the Aorist and Perfect series, case marking is on an active–inactive basis. However, in the Present series, it is on a nominative–accusative basis. In Hindi, the tripartite system is found in the Perfective aspect, while the Imperfective has a nominative–accusative system (with, as noted above, the further complication that accusative marking depends on animacy and definiteness). A particularly complex system is found in Drehu: in the Non-Past, the "agent" marker is obligatory for A and optional for S_a, i.e. the system is, in terms of our criteria, active–inactive; in the Past, this marker is used for all S and A (with some exceptions for inanimates), i.e. the system is in our terms nominative–accusative. In such cases, our general policy has been to maximize the occurrence of otherwise cross-linguistically rare types. Thus, I assign Georgian to the active type, effectively giving preference to the Aorist/Perfect over the Present; Hindi to the tripartite system, giving preference to the Perfective over the Imperfective; and Drehu to the active type, giving preference to the Non-Past.

Finally, some languages have two distinct, voice-like constructions with different case-marking alignment systems, where moreover there might be controversy as to which system should be taken as most basic. In a language with a voice system where one voice is clearly more basic than the other, as with basic Active and nonbasic

Passive in English, I take the basic voice. But some other languages, especially some Austronesian languages, are less clear. For Tukang Besi, Donohue (1999: 53) argues that the transitive construction with ergative–absolutive case marking is more basic, relative to an alternative construction with nominative–accusative case marking. (A similar problem of identifying the basic voice arises in Karo Batak (Western Malayo-Polynesian; Indonesia; Woollams 1996), but fortunately both voices have neutral case marking.) In some languages, the decisive factor is word order. For Rapanui (Oceanic; Easter Island), Du Feu (1996: 67–8) argues that the basic word order is VAP; in this order, neither A nor P is overtly case marked, and the system is neutral. However, in the alternative word orders VPA and PVA, the A requires the ergative preposition *e*, and case marking is thus ergative. In Paumarí (Arauan; Brazil; Chapman and Derbyshire 1991: 164–6), the usual word order in transitive clauses is AVP, and the case marking is ergative–absolutive; however, under alternative word orders PVA and APV, the case marking is nominative-accusative. (An added complication in Paumarí is that object pronouns can only be preverbal, so that with pronouns only the nominative-accusative system is possible.) In such cases of voice-like alternations, including those described as conditioned primarily by word order, I have followed the author of the source in deciding which alternant is basic, and have used the case-marking system of that alternant. (In the case of Tagalog (Western Malayo-Polynesian; Philippines), I took as decisive constructions in which neither S nor A nor P is in "focus", which leads to a neutral system.)

3 Map 99: Alignment of Case Marking of Pronouns

1.	Neutral	79
2.	Nominative–accusative (standard)	61
3.	Nominative–accusative (marked nominative)	3
4.	Ergative–absolutive	20
5.	Tripartite	3
6.	Active–inactive	3
7.	None	3
	total	172

Pronouns merit a separate map from full noun phrases because in many languages pronouns have a different case-marking system from full noun phrases—twenty-three languages in the sample used for these chapters (twenty-five if the marked nominative system is considered distinct from the rest of nominative–accusative). In English, for instance, while full noun phrases have a neutral case marking system, pronouns have a nominative–accusative system, e.g. with *we* for S and A but *us* for P. Some of the problems discussed in §2 apply also in the case of personal pronouns, and in such instances similar solutions are applied, with one additional criterion, namely the assignment of some weight to differentiating pronouns from full noun phrases. Thus, in English some pronouns have a nominative–accusative system, while others (in particular, second-person *you*) do not; the nominative–accusative system is selected here in part because it differs from the case-marking system for full noun phrases. Sometimes the problems that occur with full noun phrases occur with somewhat less force in the case of pronouns; for instance, several languages where full noun phrase Ps sometimes take an accusative marker and sometimes not require pronouns to take this marker (e.g. Persian), so that there is no hesitation in assigning them to the nominative–accusative type with respect to case marking of pronouns. However, pronouns bring with them a host of other problems that can lead to uncertainties in assigning languages to case-marking alignment types with respect to their pronominal systems.

These problems centre on the question of what constitutes a pronoun. In some languages, such as English, pronouns have essentially the same distribution as other noun phrases, so that there is direct comparability between pronouns and full noun phrases with respect to case-marking alignment. However, this is overall a clear minority pattern among the world's languages; in the vast majority of the world's languages it is, for instance, not necessary to have a pronoun in subject position (see Chapter 101).

For some of the languages in the sample, we have clear statements in the sources consulted that personal pronouns, at least in some of the S, A, and P positions of the clause, are simply impossible: Canela-Krahô, Wari', Wichita (see §1). For many other languages, the sources note that pronouns are rarely used in the language in question, but nonetheless the sources are reasonably clear that if pronouns are used (for instance, for contrast), then they have such-and-such a case-marking system. In yet other languages, the sources are not sufficiently explicit to enable the reader to reach a firm conclusion, with the result that several languages included in the map for full noun phrases are not included in the map for pronouns. In many of the languages covered by the present paragraph, the function of expressing the person-number of the S, A, or P of a clause is covered by person-number marking on the verb, for instance in Abkhaz (North-West Caucasian; Abkhazia, Georgia). And at least for some such languages, it has been argued that the "real" S, A, and P are not the pronouns, but rather the pronominal affixes on the verb (which fall under the topic of Chapter 100 rather than of the present chapters). However, there are also languages like Japanese which tend to avoid using pronouns but nonetheless have no encoding of the corresponding person–number information in the verb.

The real problem is caused by elements that encode such pronominal features as person-number but whose status on the scale from pronoun (as a kind of noun phrase) to pronominal marker (as a bound morpheme attached to the verb or other predicate) is unclear. In general, I have here followed the analyses of the sources, mainly because in general I have little reason to doubt these sources, although one must bear in mind that it can take a fairly sophisticated (and potentially controversial) analysis to really tease the various possibilities apart. In some instances, the analysis adopted here may be pushing against the boundaries of plausibility, for instance in treating the French clitic pronouns as relevant instances of pronouns (and thus concluding that French has a nominative–accusative case-marking system for pronouns, on the basis of clitic oppositions like nominative *je* versus accusative *me*, although the disjunctive pronoun *moi* shows no case distinction). In some instances, then, more systematic research into the nature of pronouns may require some redrawing of boundaries.

4 Geographical distribution

Some of the systems are so rare that it is questionable whether anything reliable can be said about their geographical distribution. Thus, the three instances of languages that do not allow pronouns in all of S, A, and P positions are all from the Americas, but this is surely coincidence; another such language (not in the sample) is the African language Mbay (Keegan 1997: 62–3). The three instances of the active system are more scattered, but again it is not clear that anything significant attaches to the fact that one is found in the Pyrenees (Basque), one in New Guinea (Imonda), and one in Oceania (Drehu). The tripartite system emerges primarily from the intersection of nominative–accusative and ergative–absolutive systems, and is represented by only three languages in the sample for this chapter (Hindi, Semelai, and Nez Perce).

The marked nominative has a restricted though more interesting distribution. It is found first of all in Africa, both in Afro-Asiatic languages (Middle Atlas Berber and Harar Oromo in the sample), as well as in some Nilo-Saharan languages of East Africa (Murle in the sample)—this could well be a combined genealogical–areal grouping. In a quite separate part of the world, it is found in the Yuman languages of southern California, western Arizona, and northern Baja California. Two other languages in the sample are identified as marked nominative, although it is not clear that they instantiate exactly the same phenomenon. Igbo has complex tonal interactions between subject and verb that give rise, in some TAMs, to at least the appearance of a marked subject, although the traditional analysis would speak rather of floating tones marking TAM that happen to dock onto the subject. In Aymara, the accusative is formed by deleting the final vowel of the nominative, so that although the

nominative has more phonetic material than the accusative, the direction of derivation seems rather to be from nominative to accusative (the vowel would not be predictable under the reverse derivation).

The neutral system is widespread across the world. While it might be expected in areas that otherwise have little morphology, such as South-East Asia and West Africa, it is in fact also found in languages that have complex inflectional morphologies, but where such morphology is largely confined to the verb, as for instance in Bantu languages and several languages of the Americas. The accusative system is also widespread on a global basis. Although ergative case marking is also quite widespread, it is almost completely lacking from Africa and is rare in Europe; hotbeds of ergativity include Australia and the Caucasus, to a somewhat lesser extent parts of the Americas, New Guinea, South Asia, and the Austronesian family.

5 Theoretical issues

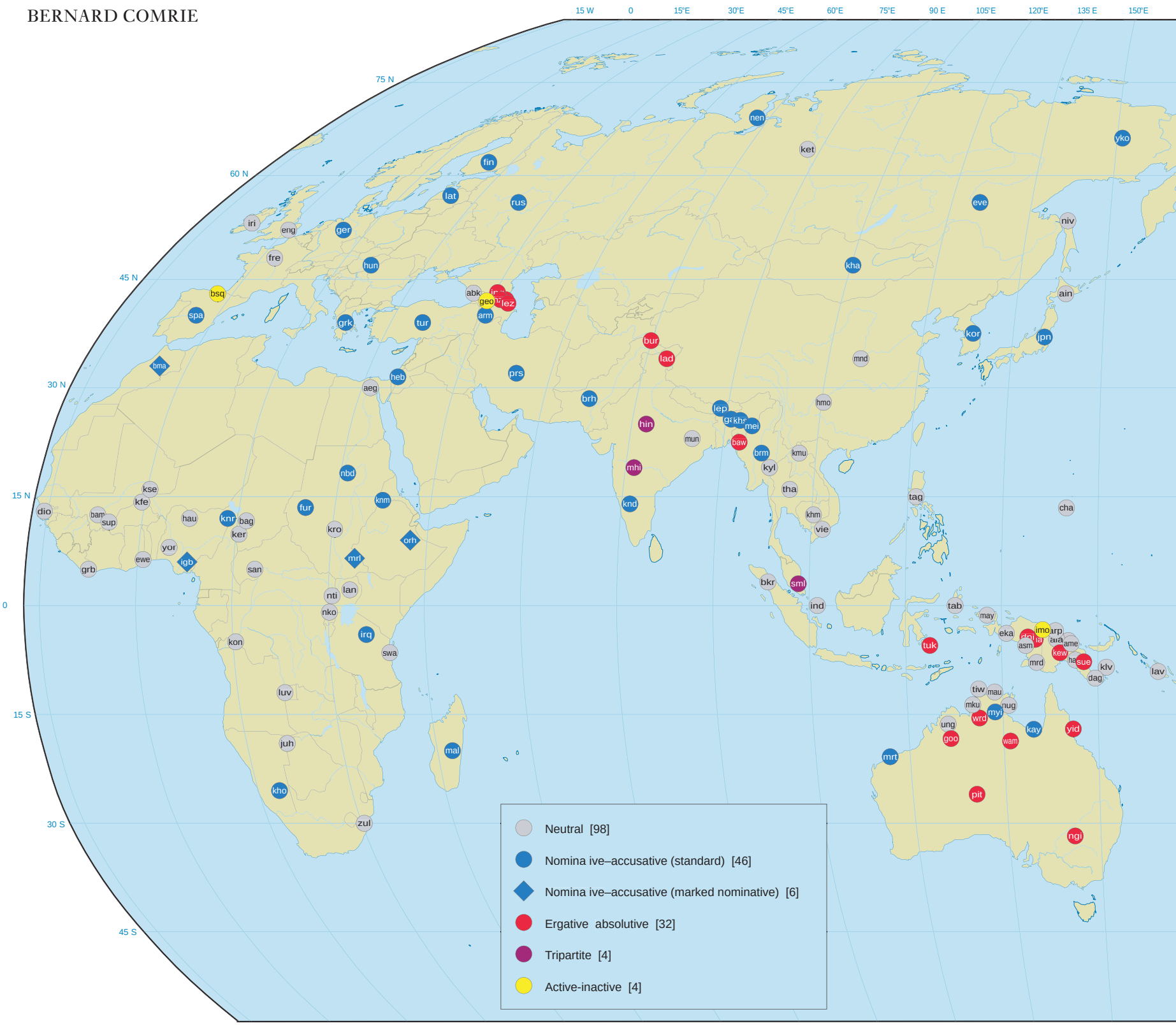
Much of the theoretical interest surrounding case-marking alignment has been concerned with correlations with other parameters, and thus features only indirectly in the present chapters. Comparison with Chapter 100 enables a comparison between the incidence of alignment patterns with noun phrases and with person markers on verbs. The scope of the *World Atlas of Language Structures* does not permit comparison with other syntactic properties, including in particular the behavioural properties of S, A, and P.

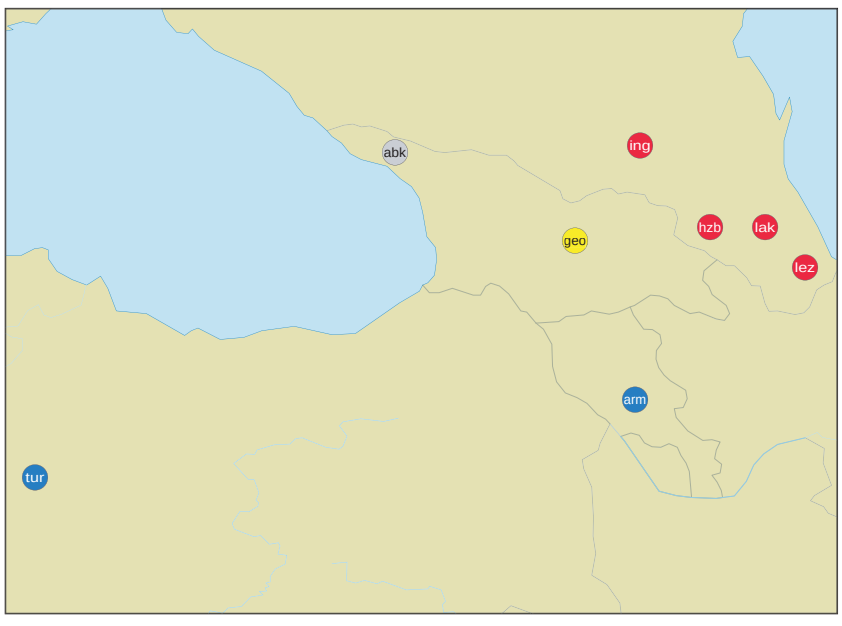
Another area of theoretical interest that goes beyond the scope of this chapter is the set of factors that determine the use of a particular case-marking pattern where the language in question allows alternatives (see, for instance, Silverstein 1976); this is because of the decision, noted in §2, to abstract away from such differences in assigning languages unequivocally to one type. Relevant generalizations would be the use of the animacy and definiteness hierarchies to constrain possible distributions of accusative and ergative case marking, with accusative case marking normally only used on a given noun-phrase type if all noun-phrase types higher on the animacy and/or definiteness hierarchies have accusative case marking; and conversely, with ergative case only used if all noun phrase types lower on these hierarchies have ergative case marking.

However, this claim can be tested on the basis of the present materials with regard to the relation between full noun phrases and personal pronouns, with personal pronouns, especially of the first and second persons, usually being claimed to be higher on the animacy hierarchy than full noun phrases, and thus more likely to have accusative case marking and less likely to have ergative case marking. This would predict that, if languages have a difference between the case-marking alignment of full noun phrases and pronouns, then the possible combinations should be (respectively) neutral (full NPs) + accusative (pronouns), ergative + accusative, and ergative + neutral, but not accusative + neutral, accusative + ergative, and neutral + ergative. And indeed this hypothesis is largely borne out by the languages forming the sample for these chapters. There are nine languages with the combination neutral + accusative, five with the combination ergative + accusative (four of them in Australia; in the fifth, Paumarí, as noted in §2, the difference correlates with different word order possibilities), and six with the pattern ergative + neutral (two of them in the Caucasus and two of them Eskimo languages). There are, however, some exceptions: two languages with the combination accusative + neutral (though one—Middle Atlas Berber—has the marked nominative subtype for nouns, i.e. a somewhat aberrant variant of nominative–accusative case marking); no languages with the combination accusative + ergative; and only one language with the combination neutral + ergative—this last language is Chamorro, which has a complex interaction of case marking of pronouns and voice-like phenomena. The Marathi combination tripartite + accusative is like ergative + accusative and ergative + neutral in having a distinctive ergative case for nouns but not for pronouns.

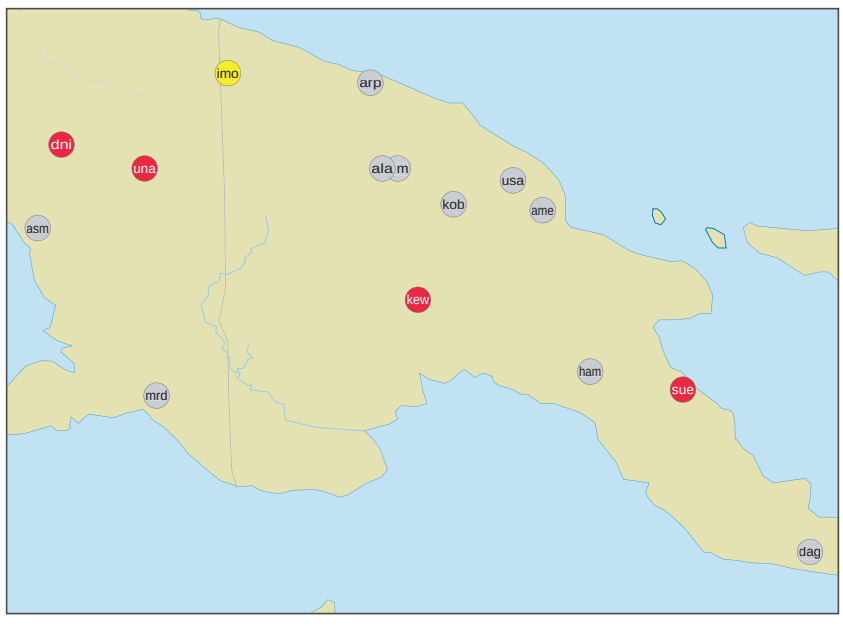
98 Alignment of Case Marking of Full Noun Phrases

BERNARD COMRIE

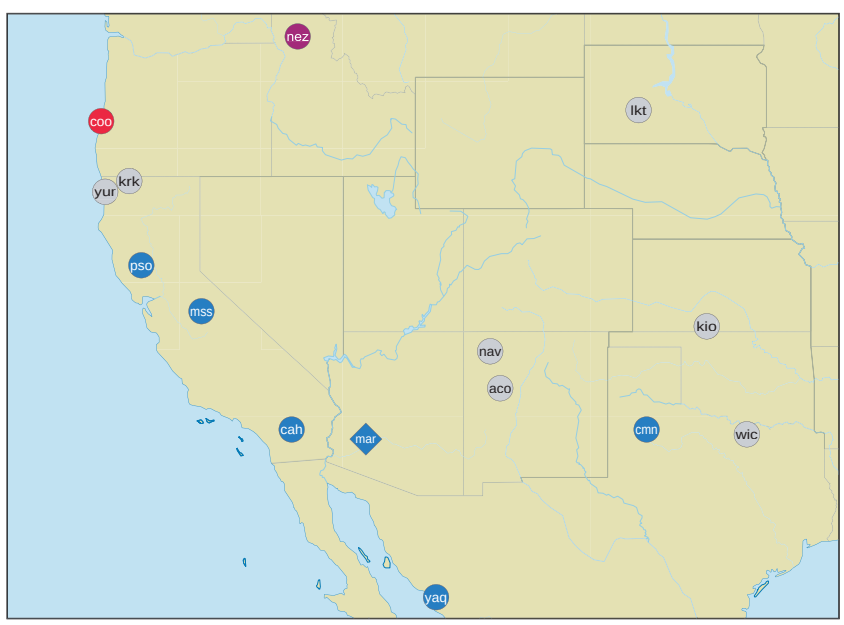




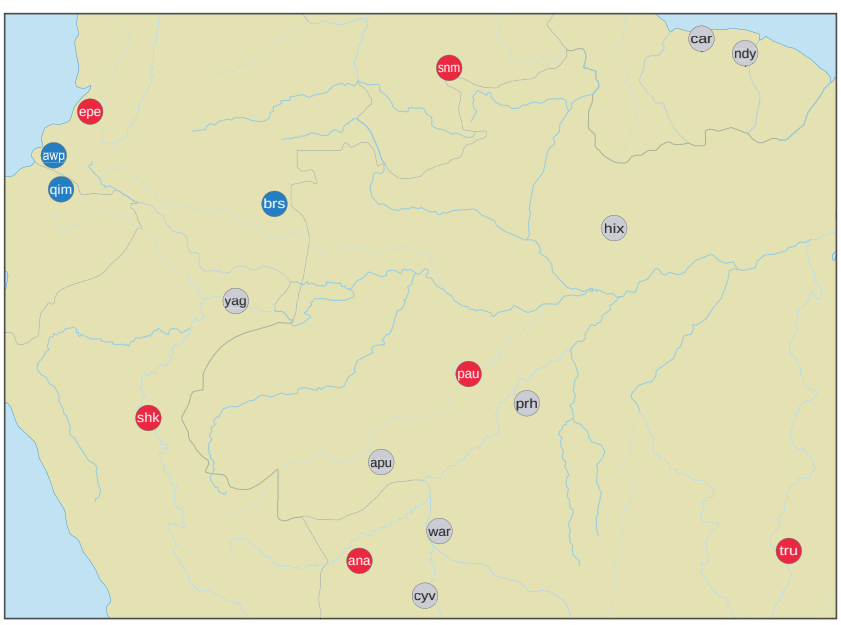
Caucasus and adjacent areas



eastern New Guinea



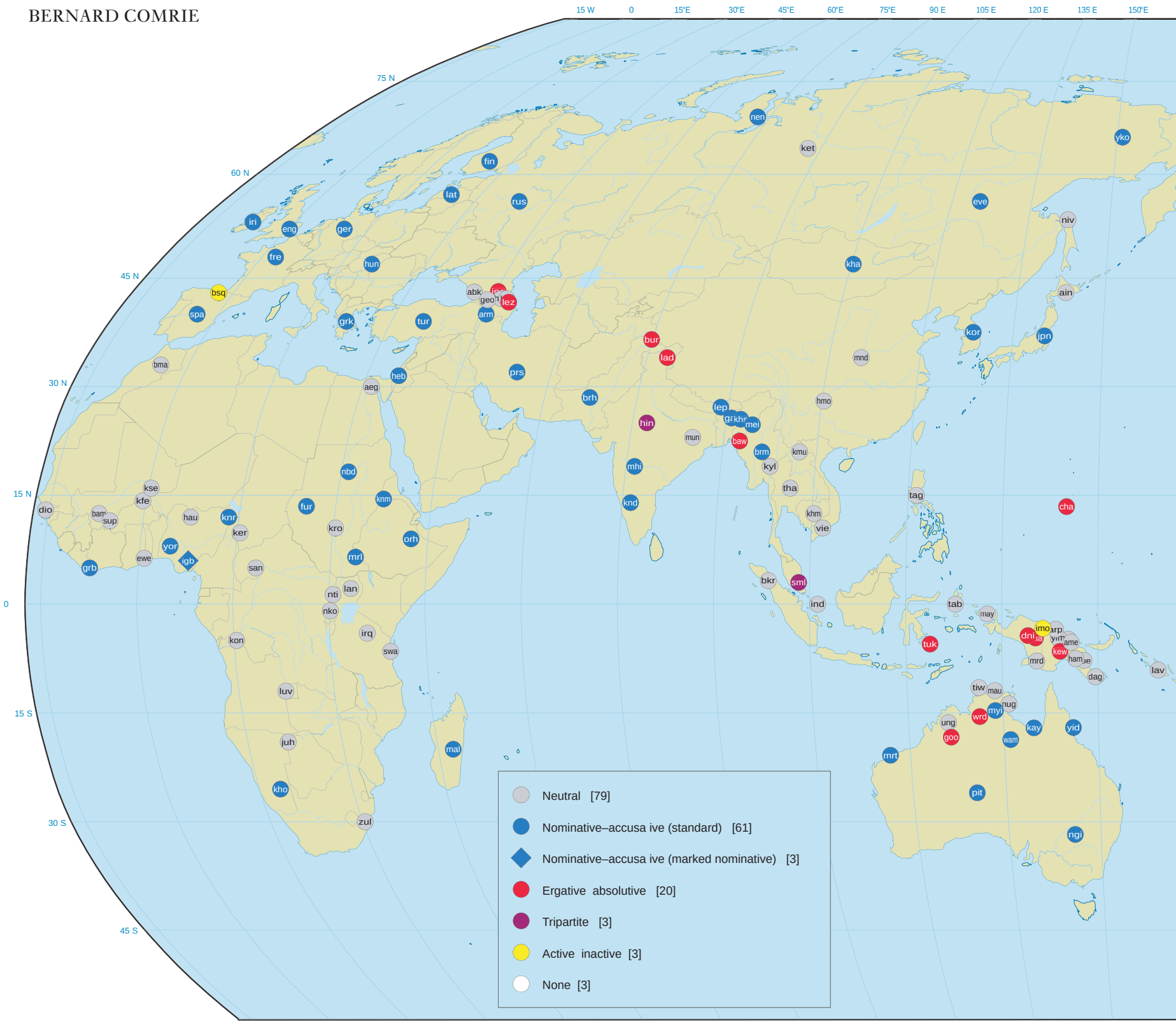
southwestern United States and northern Mexico

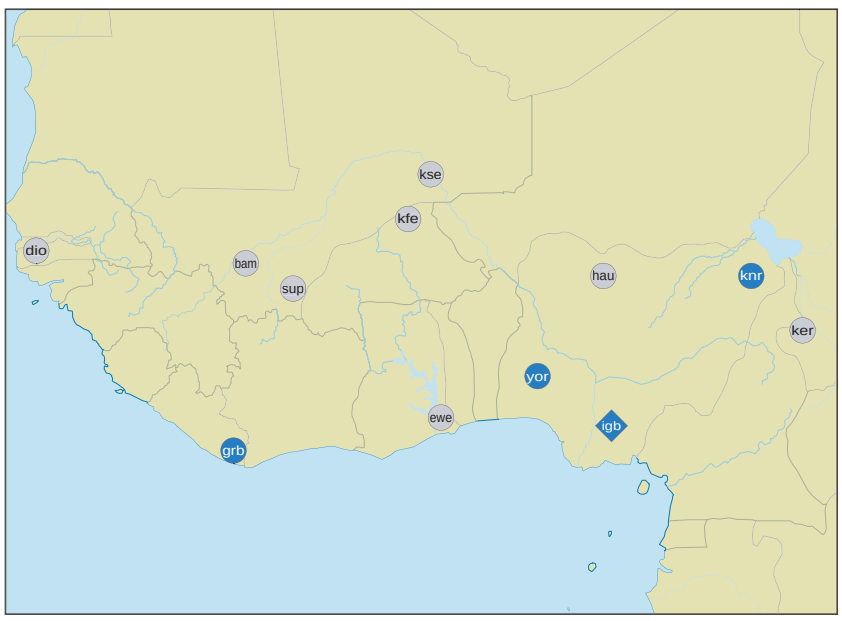


northwestern South America

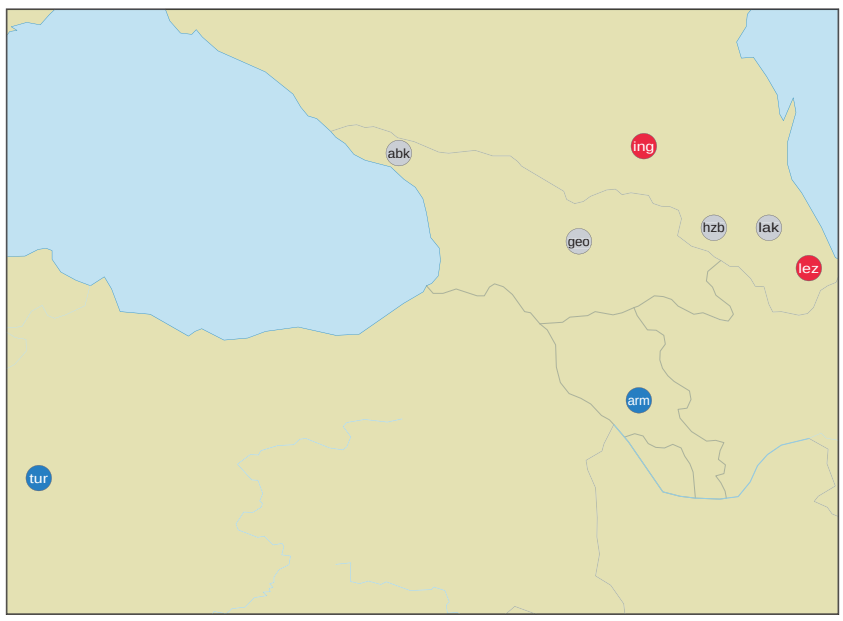
99 Alignment of Case Marking of Pronouns

BERNARD COMRIE

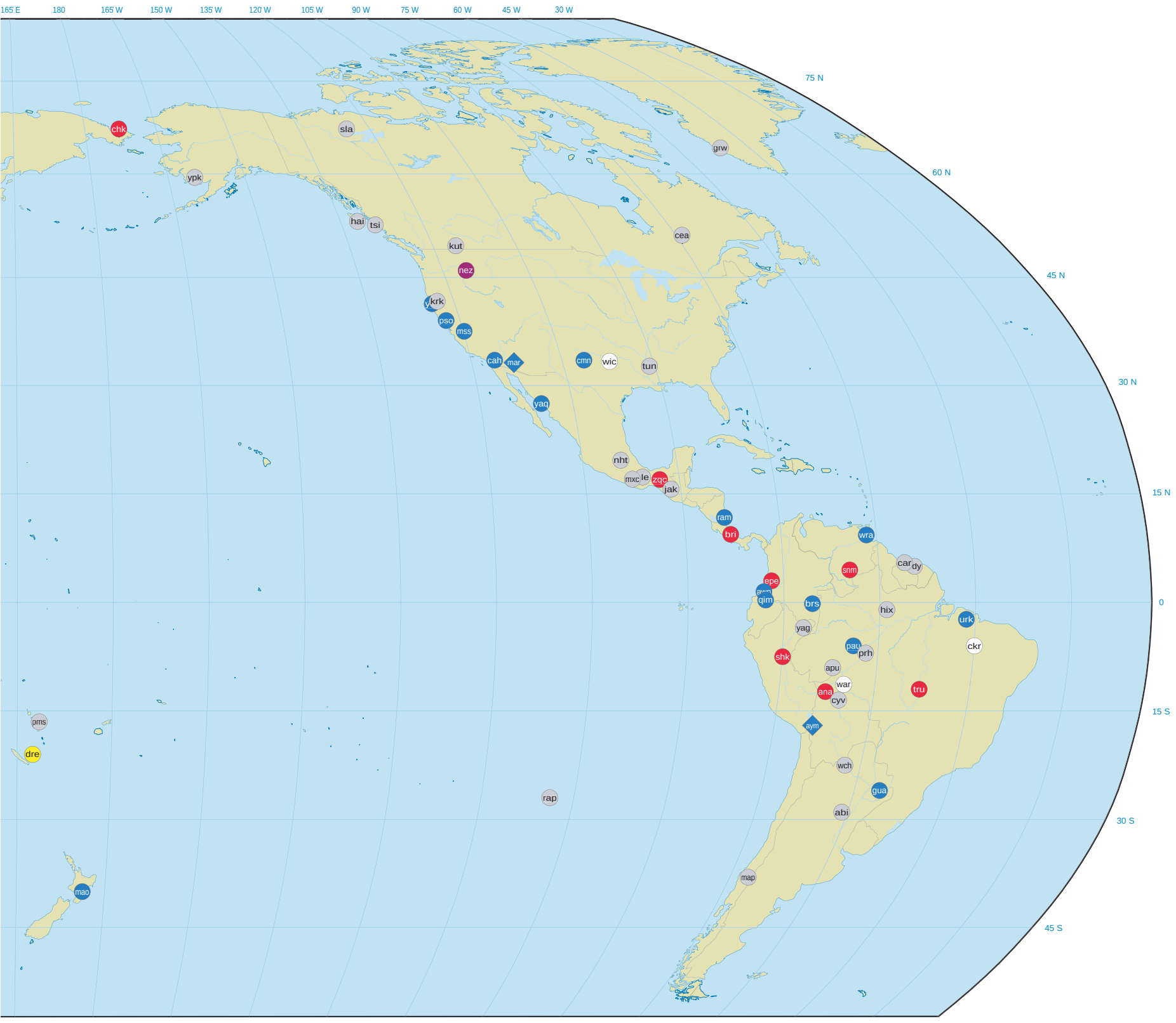


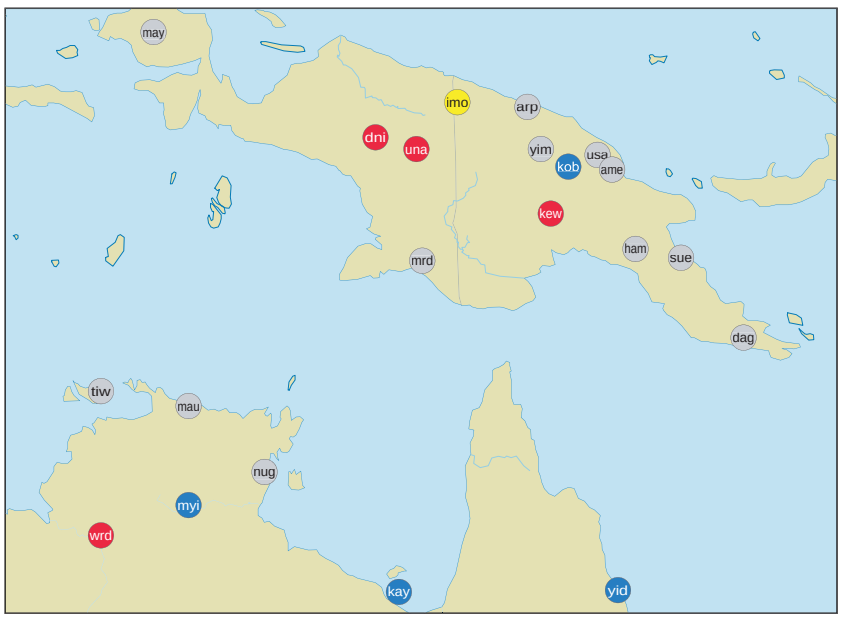


West Africa

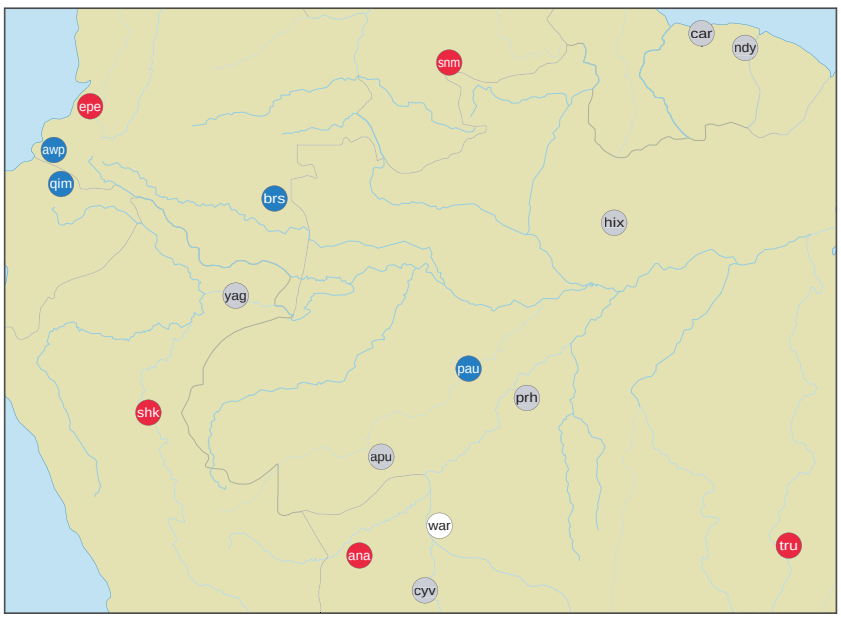


Caucasus and adjacent areas





New Guinea and northern Australia



northwestern South America

100 Alignment of Verbal Person Marking

ANNA SIEWIERSKA

1 Defining the values

Languages differ not only in regard to the number and nature of the arguments displaying verbal person marking (see Chapter 102), but also in the alignment of the person markers. The term **alignment** may be intuitively understood as reflecting how the two arguments of the transitive verb, the agentive argument (A) and the more patient-like argument (P), align with the sole argument of the intransitive verb, the S. In Map 100 the alignments of verbal person markers are represented by the following six values:

1.	Neutral alignment (no verbal person marking)	84
2.	Accusative alignment	212
3.	Ergative alignment	19
4.	Active alignment	26
5.	Hierarchical alignment	11
6.	Split alignment	28
	total	380

Neutral alignment corresponds to absence of verbal person marking.

By **accusative alignment** is understood a common treatment of the S and A and a different treatment of the P, as in (1) from Tawala (Oceanic; Papua New Guinea), where the S and A markers are prefixes and the P markers are suffixes.

(1) Tawala (Ezard 1997: 289, 116)

a. *i-bowi-ye-ya*
3SG.A-deny-TR-3SG.P
'He denied him.'

b. *apo* *i-na-nae*
FUT 3SG.S-POT-go
'He will go.'

Observe that the form of the third-person singular marker is *i-* for both the A in (1a) and the S in (1b), but *-ya* for the P in (1a).

In **ergative alignment,** by contrast, it is the S and P that are treated alike, while the A is treated in another way. This is illustrated in (2) from Konjo (Western Malayo-Polynesian; South Sulawesi).

(2) Konjo (Friberg 1996: 141, 140)

a. *na-peppe'-i* *Amir* *asung-ku*
3.A-hit-3.P Amir dog-1
'Amir hit my dog.'

b. *a'-lampa-i* *Amir*
INTR-go-3.S Amir
'Amir goes.'

Note that the third-person markers of the P in (2a) and of the S in (2b) are the same and are distinct from that of the A in (2a).

In **active alignment** there are two patterns of identification of the S: sometimes it is treated like the A and sometimes like the P, depending on a range of semantic factors such as eventhood, performance/effect, instigation, control, and significant affectedness (Mithun 1991). A language manifesting active alignment of verbal person forms is Koasati (Muskogean; Louisiana).

(3) Koasati (Kimball 1991: 189, 204, 120, 118)

a. *okolcá* *hóhca-li-halpí:s*
well dig-1SG.A-ability
'I can dig a well.'

b. *tálwa-li-mp*
sing-1SG.S-HEARSAY
'(He says) I sing.'

c. *ca-pa:-batápli-t*
1SG.P-LOC-hit-PST
'He hit me on the back.'

d. *ca-o:w-álli-laho-V*
1SG.S-LOC-die-IRR-PHR.TERM
'I will drown.'

Observe that the form of the first-person singular S marker in (3b) is the same as that of the A marker in (3a), while the form of the S in (3d) is the same as that of the P in (3c).

In **tripartite alignment** each of the three arguments S, A, and P is treated differently. Tripartite verbal person marking is rare, and when it occurs it is manifested only for part of the person paradigm. For instance, in Yukulta (Tangkic; Queensland, Australia) it is exhibited only by second-person singular and all first-person forms.

(4) Yukulta (Keen 1983: 239, 237, 215)

a. *waranaŋkuḷu-**ka**-ti*
go.NEG.DES-1SG.S-PRES
'I'm trying to go.'

b. *ṭalmata-**ŋa**-nti* *ŋita*
chop.IND-1SG.A-FUT wood
'I'll chop the wood.'

c. *ʸinkaka-**nki*** *ŋata*
follow.IMP-1SG.P me
'Follow me.'

Finally, in **hierarchical alignment** the treatment of the A and P is dependent on their relative ranking on the referential and/or ontological hierarchies. Whichever argument is the higher ranking receives special treatment, the details of which vary from language to language. The following example is from Plains Cree (Algonquian; Canada), in which person marking is determined by the hierarchy 2 > 1 > 3.

(5) Plains Cree (Wolfart 1973: 24)

a. *ki-wāpam-i-n*
2.A-see-DIR-1.P
'You see me.'

b. *ki-wāpam-iti-n*
2.P-see-INV-1.A
'I see you.'

c. *ki-wāpam-ikw-ak*
2.P-see-INV-3PL.A
'They see you.'

We see in (5) that the second person is always marked by a prefix (*ki-*) irrespective of whether its referent is in A function as in (5a) or in P function as in (5b–c). In Cree, as in many other languages which have hierarchical verbal person marking, if the higher ranking argument is a P rather than an A, a special inverse marker occurs on the verb, *iti* in (5b) and *ikw* in (5c). As we see in (5a), Cree also has a special marker for when the A outranks the P, called a *direct marker*. This is rather unusual.

In the examples of the various types of alignments given above, all three constituents, the S, A, and P, were overtly marked. However, this need not be the case. For instance, accusative alignment may involve overt marking of the S and A and no marking of the P, or, alternatively, overt marking of the P and no marking of the S and A. The same in principle holds for the other types of alignment.

While in many languages the verbal person markers always manifest one type of alignment, in others several alignments may be observed. Such **splits in alignment** may involve two or more non-neutral alignments, or a non-neutral alignment and neutral alignment. Since the term *neutral alignment* is interpreted here as meaning absence of verbal person marking, these two types of splits have been treated differently. Languages displaying splits involving two or more non-neutral alignments—for instance, accusative and ergative, or active and accusative, or ergative and tripartite—have all been classified in terms of value six, split alignment. Internal splits in alignment involving combinations of neutral and non-neutral alignment, on the other hand, have been grouped under the non-neutral alignment (thus ignoring the neutral alignment).

Both types of splits in alignment, those involving non-neutral alignments and those involving a non-neutral and a neutral alignment, may be conditioned by a range of factors. One of the factors which commonly determine alignment is person. Quite often the alignment of the first and second person differs from that of the third. For example, in many languages with active alignment, this alignment is confined to the first and second person while the third person displays neutral alignment. This is in fact the casein Koasati, mentioned above, as well as in Amuesha (Arawakan; Peru), Tlingit (Na-Dene; British Columbia), and Wichita (Caddoan; Oklahoma). The converse situation, neutral alignment of the first and second person and non-neutral of the third, is also attested, but only rarely. In Trumai (isolate; Upper Xingu, Brazil) there is ergative alignment in the third person, with S and P marked by means of a verbal enclitic *e(n)*, but no marking of either first or second person. Compare (6a–b) with (6c–d).

(6) Trumai (Guirardello 1999: 95, 99, 29)

a. *iyi* *waṭkan-e*
PCL cry-3SG.S
'She cried.'

b. *hai-ts* *ka-in* *iyi* *midoxos-e*
I-ERG PST-FOC PRT call-3SG.P
'I called him.'

c. *ha* *pita* *ka-in*
I go.out PST-FOC
'I went out.'

d. *ka'natl-ek* *ha* *midoxos*
that-ERG I call
'That one called me.'

A similar split of first and second person as opposed to third person involving neutral and non-neutral alignment, though this time accusative rather than ergative, is also found in a restricted way in English. Note that the *-s* marking of the verb in *(S)he come-s* and *(S)he like-s apples* is a manifestation of accusative verbal alignment restricted to the third person.

Person splits involving two non-neutral alignments are also to be found. One case in point is Yukulta, mentioned above, in which the first-person singular and nonsingular and the second-person singular display tripartite alignment, contrasting with the accusative alignment of the second-person nonsingular and third person.

Another common factor influencing the alignment of verbal person marking is tense and aspect. For instance, in Itzaj (Mayan; Guatemala) the verbal person markers align ergatively in the completive aspect and in dependent clauses, but accusatively in the noncompletive aspect. Still other factors which may influence alignment include mood, polarity, humanness, animacy, definiteness, word order, and main versus subordinate clause status. Typically such factors distinguish between non-neutral and neutral alignment. Thus, for example, Estonian, Manambu (Sepik; Papua New Guinea), Sentani (Sentani family; Papua, Indonesia), and Tariana (North Arawakan; Brazil) have non-neutral alignment in positive clauses but neutral in negative ones. In Ika (Chibchan; northeastern Colombia), Koegu (Surmic, Nilo-Saharan; Ethiopia), and Siuslaw (Siuslawan; Oregon) the normal non-neutral alignment contrasts with neutral alignment in imperatives. And in the following languages there is neutral alignment in relative clauses but non-neutral in main clauses: Barasano (Tucanoan; Colombia and Brazil), Gimira (Omotic, Afro-Asiatic; Ethiopia), West Greenlandic (Eskimo), Kobon (Trans-New Guinea; Papua New Guinea), Limbu (Kiranti, Tibeto-Burman; Nepal), and Turkish (Cristofaro 2003). As these examples suggest, the neutral alignment tends to occur in non-declaratives, negatives, or subordinate clauses, the non-neutral in main, positive, declarative clauses. Occasionally, however, the reverse is the case. For instance, in Nivkh (isolate; Sakhalin Island, Russia) accusative alignment is confined to imperatives. The classification of the alignments of verbal person markers depicted in Map 100 is based only on the alignments found in main, positive, declarative clauses.

2 Geographical distribution

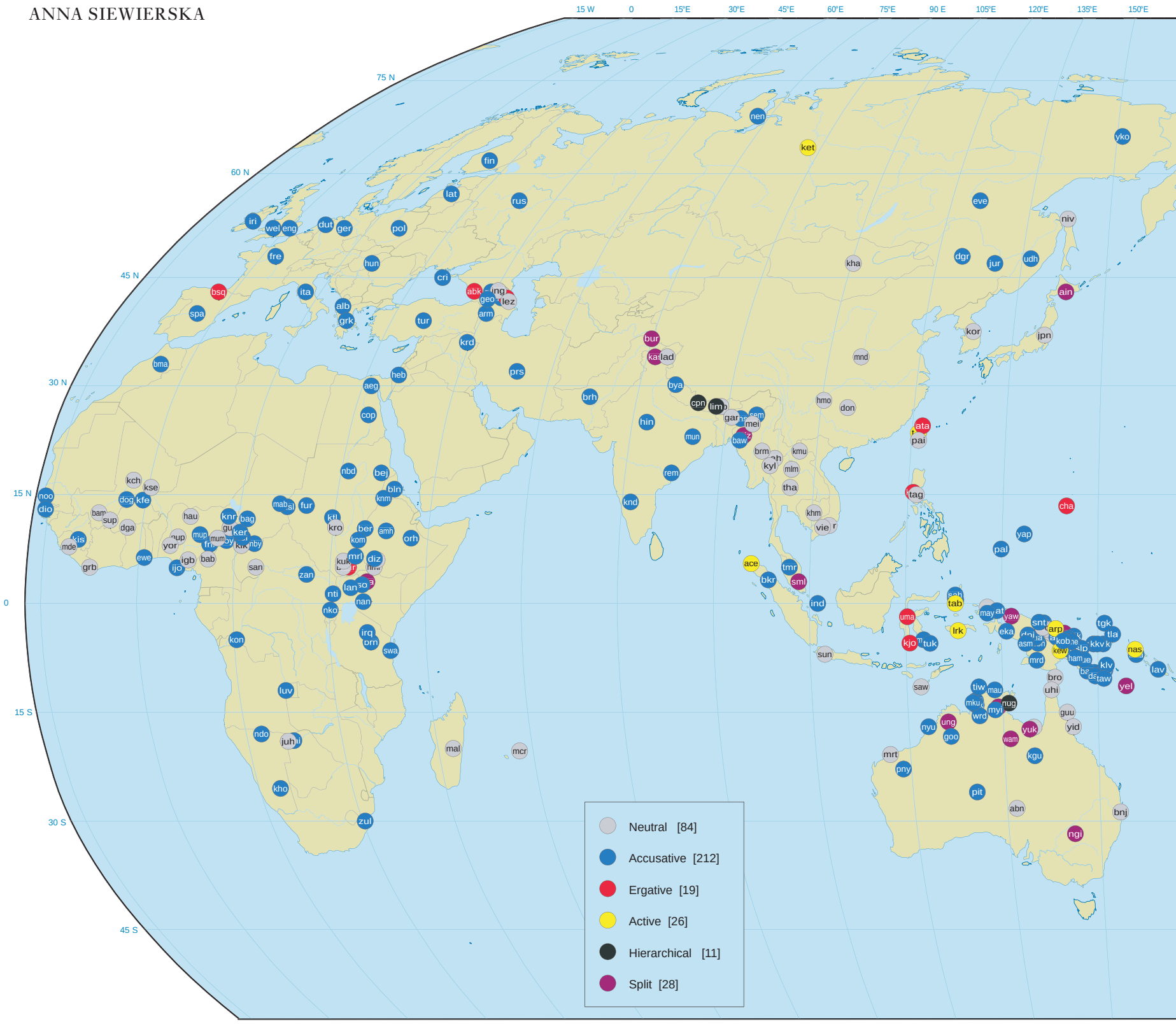
The most common alignment of verbal person markers is accusative. It occurs in around 55 per cent of the languages in the sample. It is not only the dominant alignment overall but also in every major geographical area. The next most common alignment is neutral, occurring in just over one fifth of the languages in the sample. Neutral alignment is found mainly in West Africa, the Caucasus, and South and South-East Asia. The frequency of each of the other alignment types in the sample is under 10 per cent, with split alignment being the most common (8 per cent) and hierarchical the least (3 per cent).

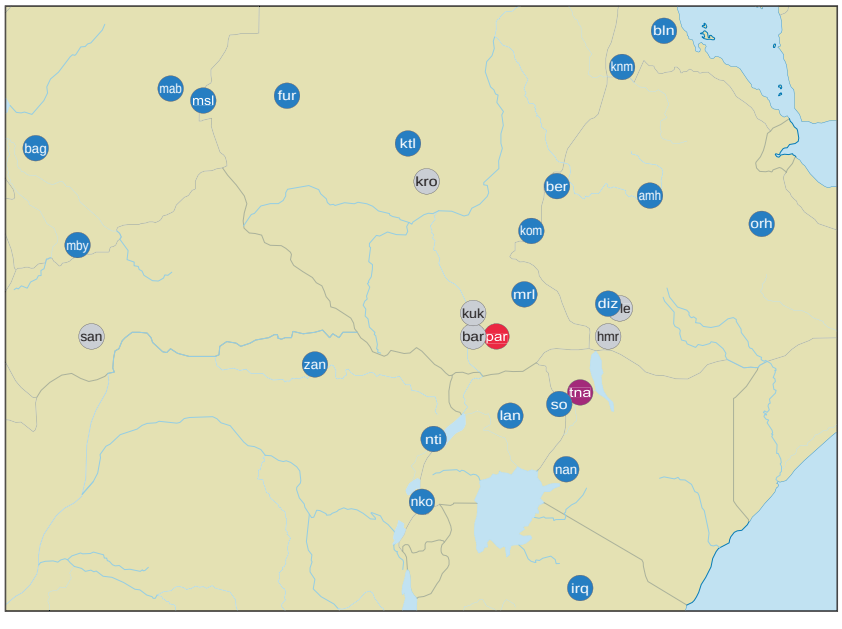
The vast majority of the split alignments involve ergative/accusative splits. Such splits are not characteristic of any particular area or areas. They are, however, more common in areas with languages manifesting some form of (unsplit) ergative alignment, be it of verbal person forms or of free pronouns (see Chapters 98–99). Ergative alignment of verbal person forms is attested mainly in South America, South-East Asia (Taiwan, Sulawesi, and the Philippines), and Mesoamerica (the Mayan languages). It is also found in Eurasia in several Caucasian languages and in Basque, as well as in the north of North America, where the Eskimo-Aleut languages are spoken. Ergative alignment of verbal person markers is absent in New Guinea, Australia, and the Pacific. One does, however, encounter sporadic instances of split accusative/ergative alignment in these areas. The area virtually devoid of ergativity is Africa. Turning to active alignment, most instances of active alignment come from the Americas. Active alignment also occurs in South-East Asia and New Guinea. It is not attested in Australia or Africa nor, apart from the isolate language Ket, in Eurasia. Hierarchical alignment, like active, is featured primarily in the languages of the Americas. The only other area in which it is in evidence, more than just sporadically, is among the Tibetan languages of India and Nepal.

The above data on the distribution of the alignment of verbal person forms correspond closely to the data presented in Nichols (1992) based on a smaller sample of 174 languages.

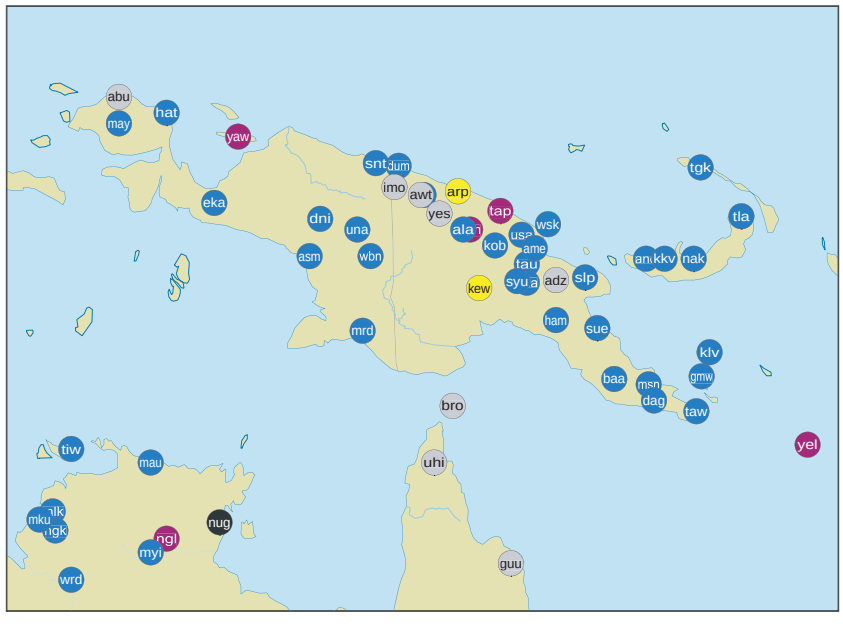
100 Alignment of Verbal Person Marking

ANNA SIEWIERSKA

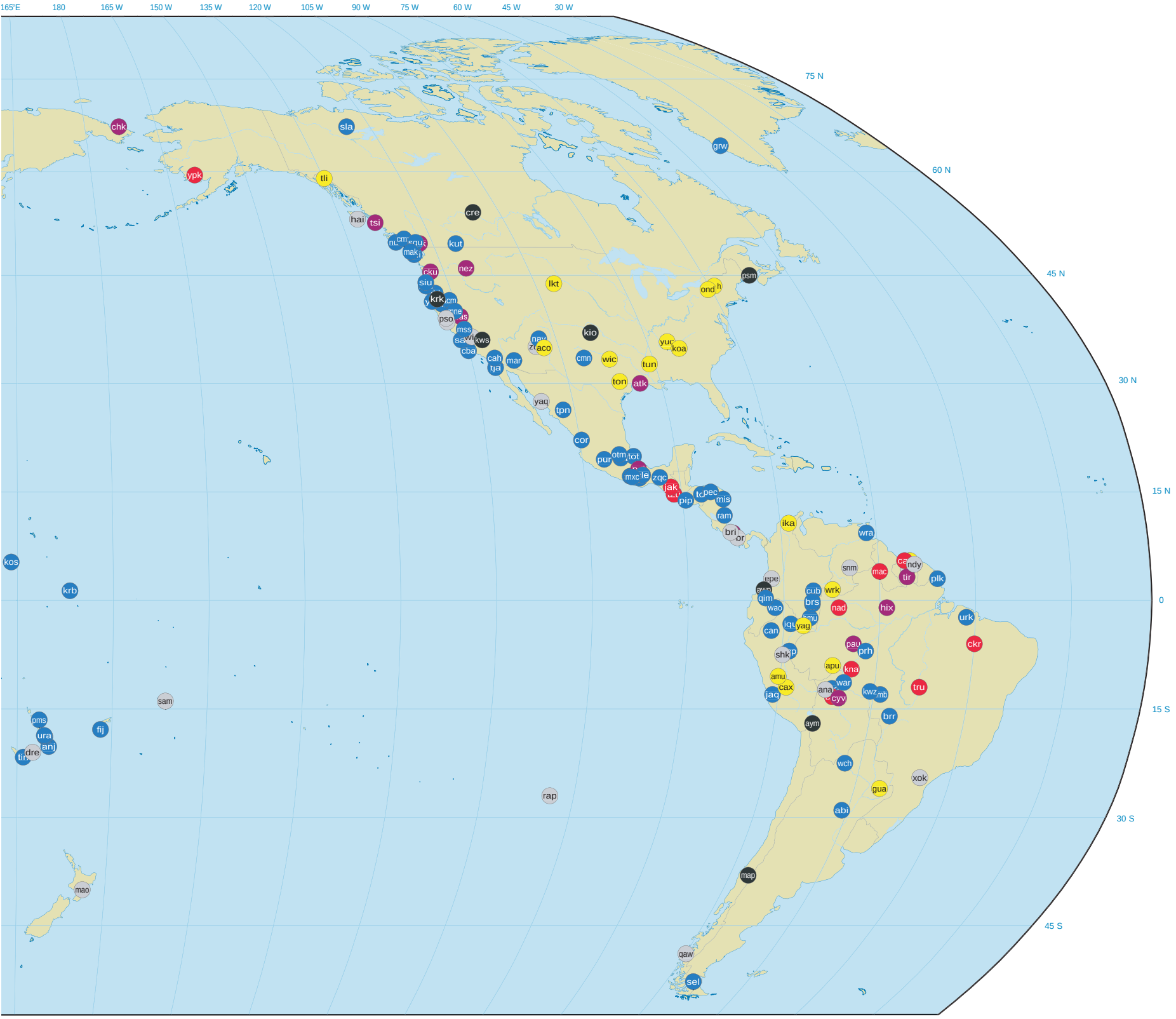


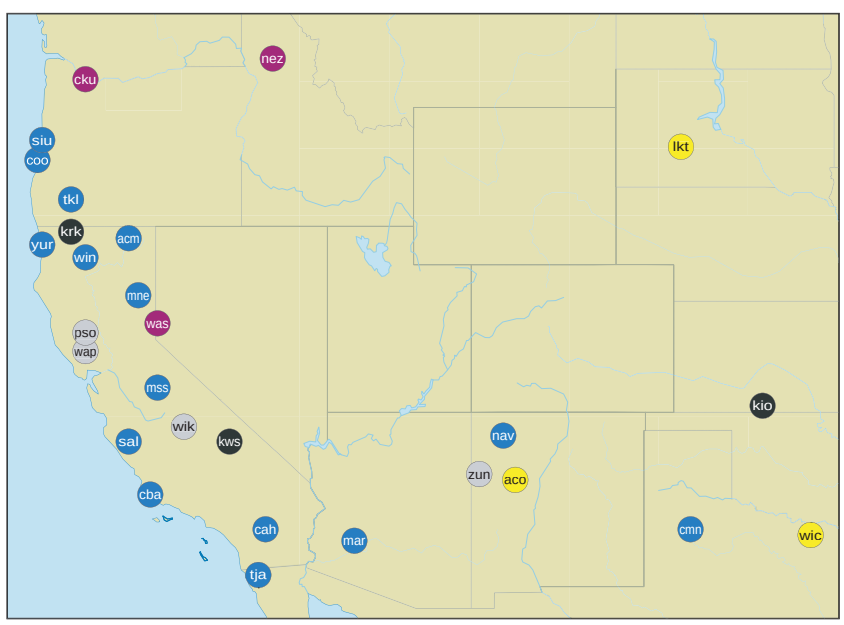


eastern Africa

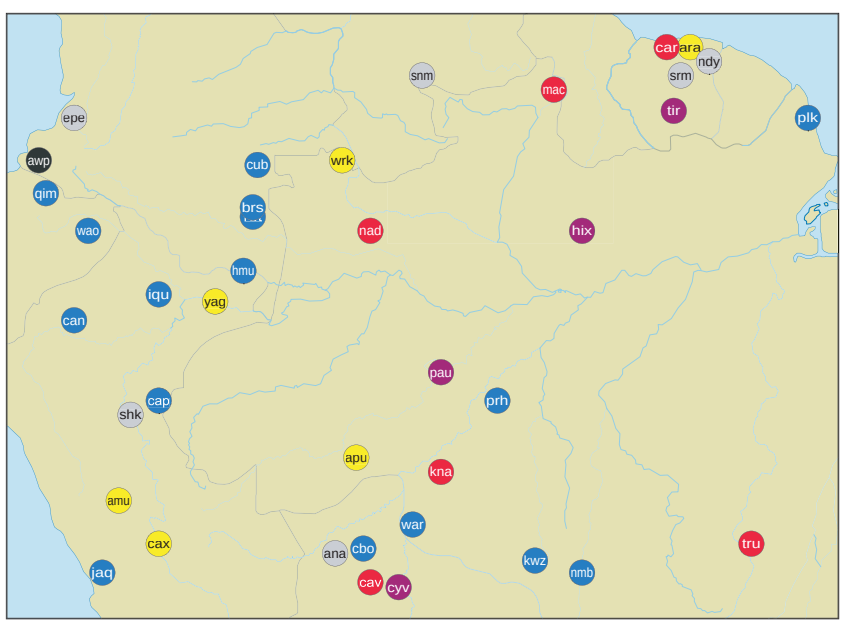


New Guinea and northern Australia





western United States



northwestern South America

101 Expression of Pronominal Subjects

MATTHEW S. DRYER

1 Defining the values

Across languages, we can distinguish two types of simple sentences differing in the nature of their subjects. One sort of simple sentence involves a **nominal subject**, a subject consisting of a noun, plus possibly some modifiers. The second sort of simple sentence involves a **pronominal subject**, where there is no nominal subject and where the subject is expressed at most by a morpheme or morphemes coding semantic or grammatical features of the subject, such as person, number, or gender. Thus, the English example in (1a) is a simple sentence with a nominal subject, while (1b) is a simple sentence with a pronominal subject.

(1) a. *The child was singing*.
b. *I was singing*.

In the English example in (1b), the morpheme expressing the subject is a separate word, an independent pronoun. In many languages, the only morpheme expressing the subject in a sentence with a pronominal subject will be an affix on the verb coding features of the subject, as illustrated by the third-person singular subject prefix *a-* in (2) from Hakha Lai (Tibeto-Burman; Myanmar); such morphemes can be called **pronominal affixes**.

(2) Hakha Lai (David Peterson, p.c.)
a-kal-tsay
3SG.SUBJ-go-PERF
'He has gone.'

For the purposes of this map, a simple sentence with a pronominal subject will be one in which the only expression of the subject is a pronominal morpheme, such as an independent pronoun or a pronominal affix on the verb. As discussed below, this also includes sentences, found in some languages, in which there is no overt morpheme at all expressing the subject, but where the sentence is interpreted in the same way as simple sentences in other languages with overt pronominal morphemes.

1.	Pronominal subjects are expressed by pronouns in subject position that are normally if not obligatorily present	77
2.	Pronominal subjects are expressed by affixes on verbs	409
3.	Pronominal subjects are expressed by clitics with variable host	33
4.	Pronominal subjects are expressed by subject pronouns that occur in a different syntactic position from nominal subjects	64
5.	Pronominal subjects are expressed only by pronouns in subject position, but these pronouns are often left out	61
6.	More than one of the above types with none dominant	30
	total	674

This map shows a number of different types of languages based on the method they use for expressing pronominal subjects. The first value shown is languages in which simple sentences with pronominal subjects **normally if not obligatorily contain a pronoun in subject position.** For the purposes of this map, subject position is defined as the position in the sentence in which nominal subjects appear. English is an example of this first type: the pronoun *I* in (1b) occurs in the same syntactic position as the nominal subject *the child* in (1a), and the pronoun is obligatory; if one leaves out the pronoun in (1b), the result is ungrammatical (**was singing*). Only declarative sentences with verbal predicates are considered; many languages that normally employ pronouns in subject position in declarative sentences do not do so in imperative sentences, for example English (e.g. *Eat your dinner*). Note that a language in which it is grammatically possible to have simple sentences without anything in subject position, but in which this option is seldom taken in actual usage, is treated as a language of the first type. For the purposes of this map, if all sentences with pronominal subjects on a couple of pages of text in a language have a pronoun in subject position, the language is coded as being of the first type.

The second type, and the one that is by far the most common (outnumbering all other types combined), is languages in which the normal expression of pronominal subjects is by means of **affixes on the verb.** Hakha Lai is an example of such a language. The example in (2) above illustrates this: the sole expression of the pronominal subject is by means of the third singular subject prefix on the verb.

The third value shown on the map is for languages in which the normal expression of pronominal subjects is by means of **pronominal clitics which have variable hosts**, attaching to different elements in the clause. For example, in Chemehuevi (Uto-Aztecan; southwestern United States), pronominal subjects are expressed by enclitics which attach to the first word in the clause. Compare (3a), illustrating SOV word order with a nominal subject, with (3b), in which the first-person singular subject clitic *=n* is attached to the first word in the sentence.

(3) Chemehuevi (Press 1979: 59, 118)
a. *Ann waha-k tɨmpi punikai-vɨ*
Ann two-OBJ stone see-PST
'Ann saw two stones.'
b. *puusi-a=n maga-vɨ*
cat-OBJ=1SG.SUBJ give-PST
'I gave a cat.'

The fourth value shown on the map is for languages where the expression of pronominal subjects is by means of **pronominal words that occur in a syntactic position distinct from that of nominal subjects**. This includes both languages where the pronoun normally co-occurs with the noun and languages where it does not. An example is given in (4a) from Longgu (Oceanic; Solomon Islands), where there is a separate pronominal word *e* expressing the fact that the subject is third person singular. However, this pronominal word is not in subject position, as can be seen in (4b), where the same pronominal word *e* co-occurs with a separate nominal subject.

(4) Longgu (Hill 2002: 548, 550)
a. *e zudu*
3SG sit
'He/she is sitting.'
b. *mwela-geni e vusi angi*
child-woman 3SG almost cry
'The girl is almost crying.'

The fact that the pronominal word *e* is obligatory, even when there is a separate nominal subject, means that it is in a distinct syntactic position from that of the subject. In fact, this word is somewhat analogous to the subject prefix in (2) from Hakha Lai, except that it is a separate phonological word. Such words are probably best viewed as being part of a verbal complex, which contains the subject pronominal word and the verb, plus various other grammatical words associated with the verb.

Also included as instances of the fourth type are languages which are like Longgu in having a word expressing pronominal features of the subject, but where these words also code tense-aspect-mood features of the clause. For example, in Gela, like Longgu an Oceanic language of the Solomon Islands, there are words preceding the verb that code both pronominal features of the subject and tense, as illustrated in (5).

(5) Gela (Crowley 2002: 532)

a. *e* *vaŋa* *kake*
3SG.PST eat taro
'He/she ate taro.'

b. *ku* *riɣi-a* *na* *kau*
1SG.FUT see-3SG.OBJ ART dog
'I will see the dog.'

The third and fourth types are like languages with pronominal subject affixes on verbs, in that there are pronominal morphemes somewhere in the clause, though in a position distinct from nominal subjects. The fifth type, in contrast, involves languages in which **the only expression of pronominal subjects involves pronouns in subject position, but in which such pronouns are optional**, so that if they are absent, there will be no expression of the subject at all. An example of such a language is Japanese, as in (6).

(6) Japanese (Hiroko Nakagawa, p.c.)
tegami *o* *yon-da*
letter OBJ read-PST
'I/we/he/she/they read the letter.'

In languages where this is possible, the pronoun will typically only be left out if the reference is clear in the context. Such sentences will be considered here as instances of sentences with pronominal subjects, in the sense that they will always be assigned an interpretation that is the same as that of a sentence containing an overt pronoun —or, if the language does not allow independent pronouns in subject position, the same interpretation as a simple sentence in another language that does allow such pronouns. In describing such sentences here as having pronominal subjects, no commitment is intended as to the appropriate grammatical analysis of such sentences; in particular, this is not meant to imply that there is some level of grammatical structure which contains an element that might be described as pronominal. Rather, the classification is based entirely on the interpretation of sentences like (6) above.

The final value shown on the map is for languages which involve a mixture of the first five types. This includes languages in which **certain types of simple clauses require a pronoun in subject position, while other types of simple clauses do not**. For example, in Jakaltek (Mayan; Guatemala), third-person pronouns (which code noun class) are normally obligatory, while first- and second-person pronouns are normally absent (Craig 1977). Kenga (Central Sudanic; Chad) and Finnish are similar to Jakaltek (Vandame 1968: 35; Sulkala and Karjalainen 1992: 120, 272). Dinka (Nilotic; Sudan) and Lamani (Indic) appear to be the opposite of these languages: third-person pronouns can be left out, but first- and second-person pronouns are normally employed (Nebel 1948: 53; Trail 1970). In Nias (Austronesian; western Indonesia), obligatoriness depends on whether the clause is transitive and on mood. In intransitive realis clauses, independent pronouns are normally required in subject position unless the subject is third person inanimate; in transitive clauses and in all irrealis clauses, pronouns are not required (Brown 2001). This distribution is clearly related to the fact that the language has pronominal prefixes on the verb for transitive subjects in realis clauses and for all subjects in irrealis clauses.

This type also includes languages which involve a mixture of the second, third, and/or fourth types. For example, Chalcatongo Mixtec (Mexico; Macaulay 1996) has a combination of subject clitics and pronominal subject words that occur in a position distinct from subject position. Languages in which pronouns in subject position are obligatory or normal, but which also have pronominal morphemes of one of the second, third, or fourth types, are treated as being of the first type (rather than of the mixed type). English is an instance of this, since there is at least some coding of pronominal features of the subject on verbs (*I walk* versus *he walks*, *I am* versus *you are*).

2 Geographical distribution

Languages in which pronominal subjects are expressed by pronominal affixes are widespread throughout the world. There are two areas where this type is absent or almost absent. One is an area in South-East Asia stretching from Myanmar to Vietnam. The other is Australia outside of the north-central region, largely coinciding with Pama-Nyungan languages.

The remaining types are all considerably less frequent, but exhibit striking geographical patterns. Northern Europe and West Africa are two areas in which languages requiring or normally using pronouns in subject position are particularly common. Most of the languages shown with subject clitics are in North America or Australia. Most of the languages with pronominal words not in subject position are in sub-Saharan Africa or the western Pacific. Most of the languages that only have subject pronouns, but where it is common not to use them, are in East and South-East Asia or in Australia.

3 Theoretical significance

The distinction between languages in which subject pronouns are required and those in which they are not is sometimes described in terms of a notion of "Pro Drop" (see, e.g., Chomsky 1981: 240–8). Under some implementations of this idea, for at least some languages, it is claimed that there is underlyingly a pronoun in subject position that is deleted in sentences lacking a subject pronoun. However, given the fact that languages that require a pronoun in subject position are fairly infrequent, there is a danger of such an approach being Anglocentric, analysing other languages as being underlyingly like English despite their superficial differences. An alternative view (e.g. Van Valin and La Polla 1997: 331) that has been proposed for languages where pronominal affixes on verbs are the normal expression of pronominal subjects is that the affixes themselves are the real subjects of the clause, even in clauses in which there is a separate nominal subject; under this view the separate nominals are not really subjects, but noun phrases in apposition to the pronominal affix on the verb. Note that in characterizing languages like Hakha Lai here as not requiring a pronoun in subject position, we intend the notion 'subject position' to cover such independent nominals.

101 Expression of Pronominal Subjects

MATTHEW S. DRYER

Obligatory pronouns in subject position [77]

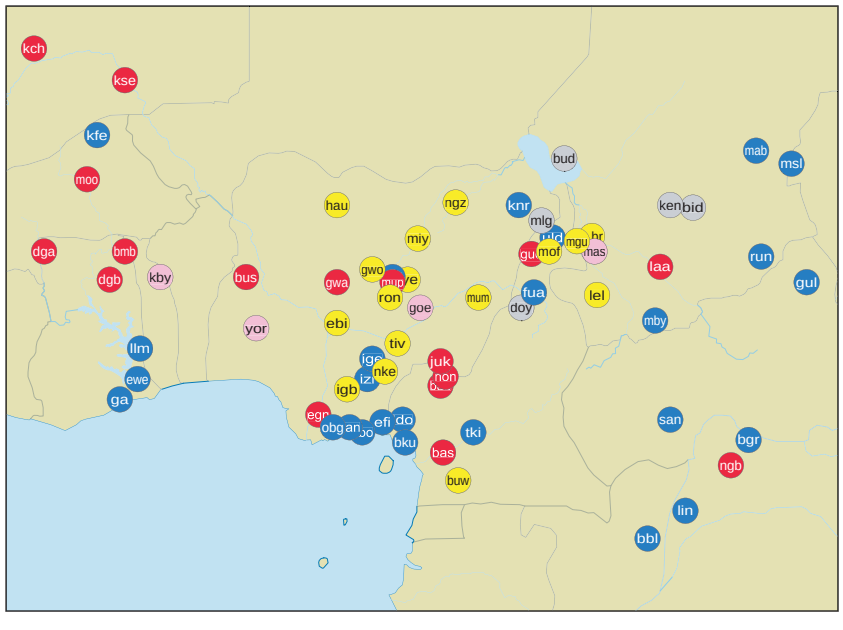
Subject affixes on verb [409]

Subject clitics on variable host [33]

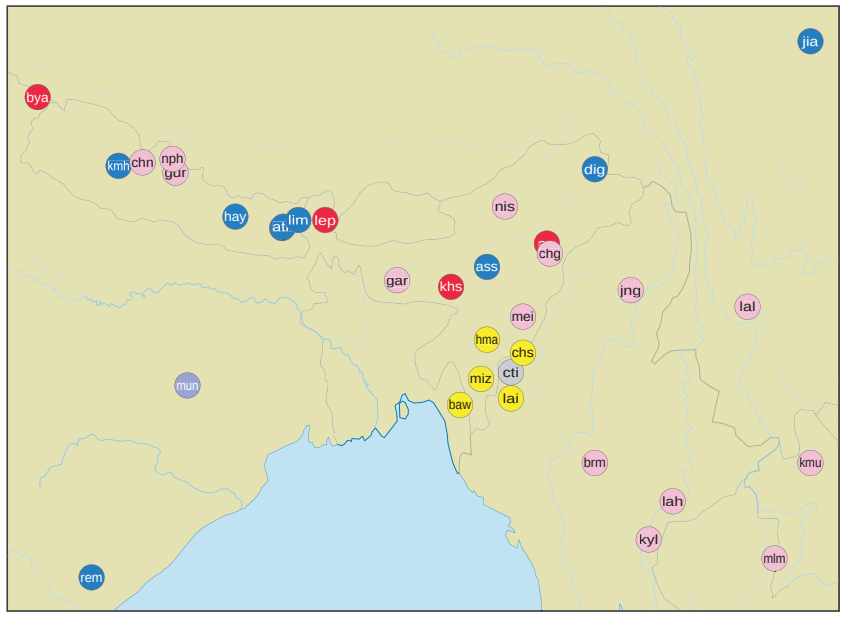
Subject pronouns in different position [64]

Optional pronouns in subject position [61]

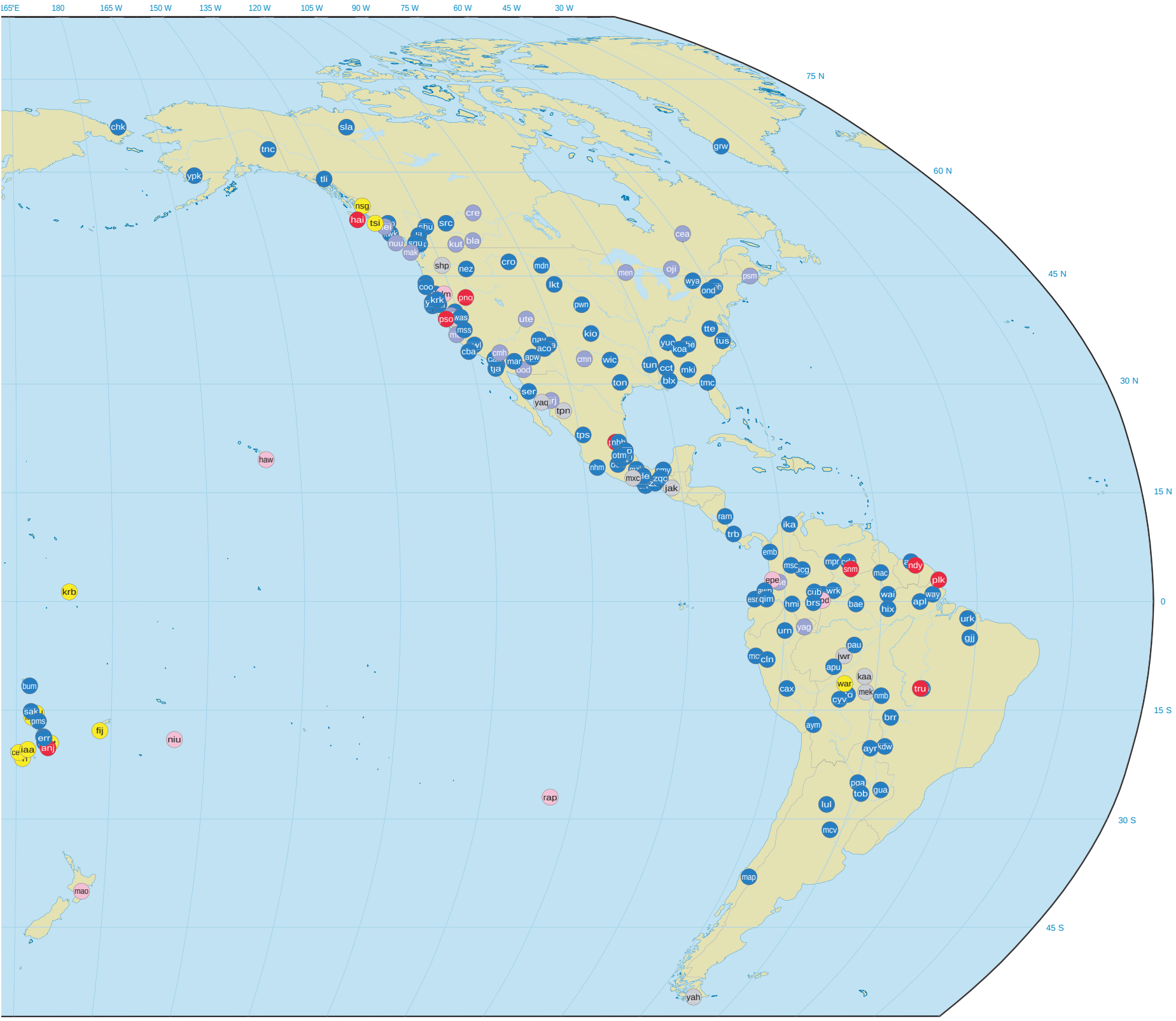
Mixed [30]

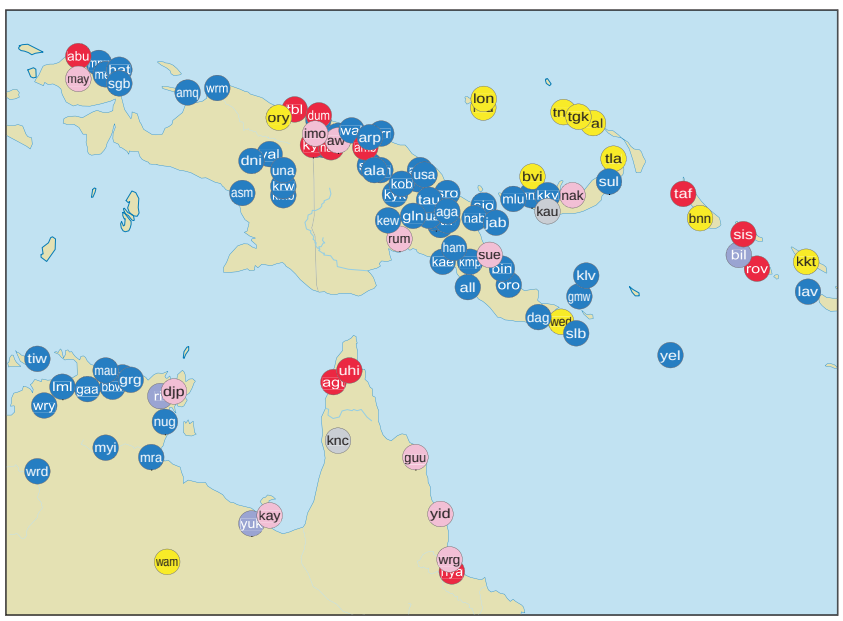


Nigeria and surrounding areas

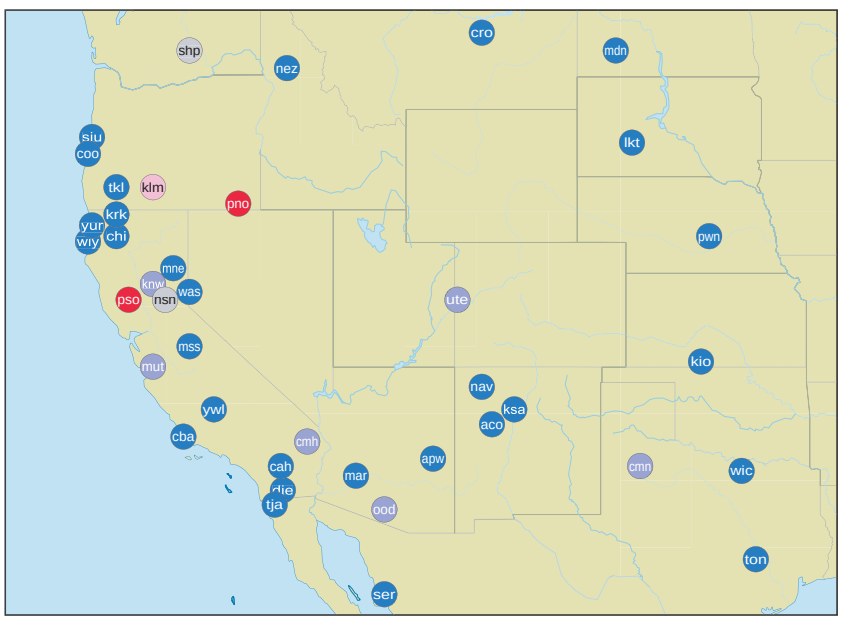


Nepal to Myanmar





New Guinea and northern Australia



western United States

102 Verbal Person Marking

ANNA SIEWIERSKA

1 Defining the values

Map 102 depicts the number and identity of the arguments of a transitive clause which display person marking on the verb. "A" stands for the agentive argument and "P" for the patient argument. Five feature values are represented:

○	1. No person marking of any argument	82
●	2. Person marking of only the A argument	73
●	3. Person marking of only the P argument	24
●	4. Person marking of the A or P argument	6
●	5. Person marking of both the A and P arguments	193
	total	378

The **lack of verbal person marking** is illustrated in (1) by Bari (Nilotic, Nilo-Saharan; Sudan).

(1) Bari (Spagnolo 1933: 17)
nan *nyanyar* *yaŋgo'*
I love mother
'I love my mother.'

Person marking of only the A argument is exemplified in (2) from Kobon (Trans-New Guinea; Papua New Guinea).

(2) Kobon (Davies 1981: 185)
yad *kaj* *pak-nab-in*
I pig strike-FUT-1SG.A
'I will kill a pig.'

An example of **person marking of only the P argument** is given in (3) from Yawa (isolate; Papua, Indonesia).

(3) Yawa (Jones 1986: 40)
Dorpinus *po* *Marianna* *r-anepata*
Dorpinus ERG.3SG.M Marianna 3SG.F.P-hit
'Dorpinus is/was hitting Marianne.'

Person marking of the A or P but not of both is found, for instance, in Nocte (Naga, Tibeto-Burman; India).

(4) Nocte (Das Gupta 1971: 21)
a. *nga-ma* *ate* *he(i)tho-ang*
I-ERG he.ACC teach-1SG.A
'I will teach him.'
b. *ate-ma* *nga-nang* *he(i)tho-h-ang*
he-ERG I-ACC teach-INV-1SG.P
'He will teach me.'

Observe that the person marking on the verb in (4a) is of the A, while in (4b) it is of the P. Whether it is the A or the P that is marked on the verb is dependent on which is higher on the person hierarchy 1 > 2 > 3. If the higher ranking argument is a P rather than an A, the verb occurs with a special marker, *h*, in (4b), called an *inverse marker*.

The final value, **person marking of both the A and P arguments**, is shown in (5) from Tawala (Oceanic; Papua New Guinea).

(5) Tawala (Ezard 1997: 99)
kedewa *kamkam* *i-uni-hi*
dog chicken 3SG.A-kill-3PL.P
'A dog killed the chickens.'

In most languages exhibiting verbal person marking, the person markers are **affixes**, attached to the lexical or auxiliary verb, as in the examples above. There are, however, also languages in which the person markers are not affixes but **clitics**. Clitics resemble affixes in forming a phonological unit with a word (their host), preceding the host (proclitics) or following it (enclitics). However, whereas affixes attach only to specific types of words (or stems), clitics attach to phrases and/or specialized syntactic positions. A common position for person clitics is the beginning of the verb phrase or verb complex. In Tinrin (Oceanic; New Caledonia), for example, the subject person markers are proclitic to the first element of the VP, which may be the verb (6a), a tense-aspect particle (6b), or a verbal modifier (6c).

(6) Tinrin (Osumi 1995: 215, 177, 182)
a. *treanrü* ***rri=fi*** *winrô* *fònri*
people 3PL.S=go follow river
'People went along the river.'
b. ***u=nrâ*** *fi* *wai*
1SG.S=PROG go already
'I am going now.'
c. ***rri=see*** *saafi* *wake* *nyôrrô*
3PL.S=NEG all.together always cook
'They do not always cook together.'

Another common clitic position is after the first word or constituent in the utterance, as is the case in many southern Uto-Aztecan languages, including Southeastern Tepehuan (Uto-Aztecan; Durango, Mexico). We see below that the person marker is encliticized to a verb in (7a), to the end of a postpositional phrase in (7b), and to a question word in (7c).

(7) Southeastern Tepehuan (Willett 1991: 194, 191)
a. *vacócoi=m=ɨt* *gu* *a'ahl*
sleep=3PL.S=PERF ART children
'The children have gone to sleep.'
b. *aptuvús* *ta'm=ach=ich* *vají*
bus on=1PL.S=PERF went
'We went by bus.'
c. *pá=pim* *duc* *va-'aiy-a'* *mu-ja'p*
when=2PL.S PCL RLZ-arrive-FUT there-area
jam-quiquia'am
2PL-homes
'When will you arrive at your homes?'

Though strictly speaking person markers such as those in Tinrin and Southeastern Tepehuan are not verbal person markers, they are treated here as such, provided that the verb is one of the constituents to which the person clitics may attach. By contrast, person markers which cannot be bound to the verb, i.e. independent person forms such as the subject markers in Woleaian (Oceanic; Micronesia) in (8), or free-standing combinations of person forms fused with tense as in Anejom (Central-Eastern Oceanic; Vanuatu) in (9), have been excluded.

(8) Woleaian (Sohn 1975: 93–4)
sar *kelaal* ***re*** *sa* *tangileng*
those children 3PL.S ASP cried
'Those children over there cried.'

(9) Anejom (Lynch 1982: 119)
et *aviñ* *numu* *aen*
3SG.AOR.A want fish he
'He wants a fish.'

Apart from affixal or clitic person marking, we occasionally find **person marking via phonological changes** in the stem. For example, in Misantla Totonac (Totonacan; Veracruz, Mexico), the second-person singular A (and also S) marker is (with some verbs) marked by suppletion of the stem. This is the case with all verbs

containing the roots /min/ 'come' and /an/ 'go'. An example of the latter is given in (10).

(10) Misantla Totonac (MacKay 1999: 153)

a. *kit* *ʔik-án*
I 1SG.S-go
'I go.'

b. *wíš* *pín*
you 2SG.S.go
'You go.'

c. *ʔút* *Ø-ʔán*
(s)he 3SG.S-go
'(S)he goes.'

Such forms of person marking have also been taken into account here.

Languages differ considerably in regard to the **conditions** under which they display verbal person marking. In some, verbal person markers are obligatory. In others they occur only in the absence of free nominal and pronominal arguments. As shown in (11), the latter is the case in Macushi (Carib; Brazil, etc.) with respect to the person markers of both the A and the P.

(11) Macushi (Abbott 1991: 24)

a. *t-ekîn* *era'ma-'pî* *paaka* *esa-'ya*
REFL-pet.ABS see-PST cow owner-ERG
'The owner of the cow saw his own pet.'

b. *i-koneka-'pî-i-ya*
3SG.P-make-PST-3SG.A-ERG
'He made it.'

Rather less commonly, person markers occur in the presence of free pronouns but not nouns (e.g. Welsh, Luo) or alternatively in the presence of nouns but not free pronouns (e.g. Palauan). The occurrence of verbal person markers may also be conditioned by the location of the NP argument corresponding to the person marker (e.g. Apurinã, Cora, Paumarí, Yagua), its animacy, definiteness, or referential status (e.g. Daga, Lower Grand Valley Dani, Hua, Persian, Usan), the particular person involved (first, second, or third) (e.g. Imbabura Quechua, Koasati, Panyjima, Parecis), as well as tense (e.g. Wolaytta), aspect (e.g. Awa Pit), mood (Evenki, Ika, Arapesh), and polarity (e.g. Estonian, Sentani).

In classifying languages in terms of the number and identity of the arguments of a transitive clause manifesting person marking on the verb, most of the restrictions mentioned above have not been taken into account. Thus languages have been classified as exhibiting A, P, or both A and P person marking on the basis of the maximal person marking that can occur in a single clause. Person markers which have not been included are: those found solely in dependent as opposed to independent clauses (e.g. Krongo), those occurring only in a minor mood such as the imperative (e.g. Nivkh), and those accompanying only a handful of verbs (e.g. Ingush).

It must be pointed out that some languages may lack markers of person on the verb for the A or the P, or even for both, but may have markers of number and/or gender. For instance, as shown in (12), in Archi (Nakh-Daghestanian; eastern Caucasus) there are prefixes on the verb (lexical or auxiliary) indicating the gender and number of the P (and also of the S) but not person.

(12) Archi (Kibrik 1972: 124)

a. *wez* *dija* *ǩ'anši* *w-i*
1SG.M.DAT father like M.P-is
'I like father.'

b. *dez* *buwa* *ǩ'anši* *d-i*
1SG.F.DAT mother like F.P-is
'I like mother.'

c. *wez* *un* *ǩ'anši* *w-i*
1SG M.DAT you like M.P-is
'I like you (male addressee).'

d. *was* *zon* *ǩ'anši* *w-i*
YOU.DAT me like M.P-is
'You like me (male speaker).'

That the verbal prefixes indeed mark only gender and not person is evidenced by the fact that while the prefix *w-* is used when the P is of male rational gender (12a) and the prefix *d-* when the P is of female rational gender (12b), the same *w-* prefix is used in (12a), (12c), and (12d), though the person of the P changes from third to second to first. To give another example, in Khanty (Ob-Ugrian, Uralic; Siberia) as well as various other Uralic languages (e.g. Hungarian, Nenets, Mordvin, Mansi, and Selkup) there are person suffixes for the A which are preceded by suffixes for the P. The latter, however, indicate only number and not person, as shown in (13).

(13) Khanty (Nikolaeva 1999: 24)

a. *we:rlə-ɲil-an*
make-DU.P-2SG.A
'You make us/them two.'

b. *we:rlə-l-an*
make-PL.P-2SG.A
'You make us/them.'

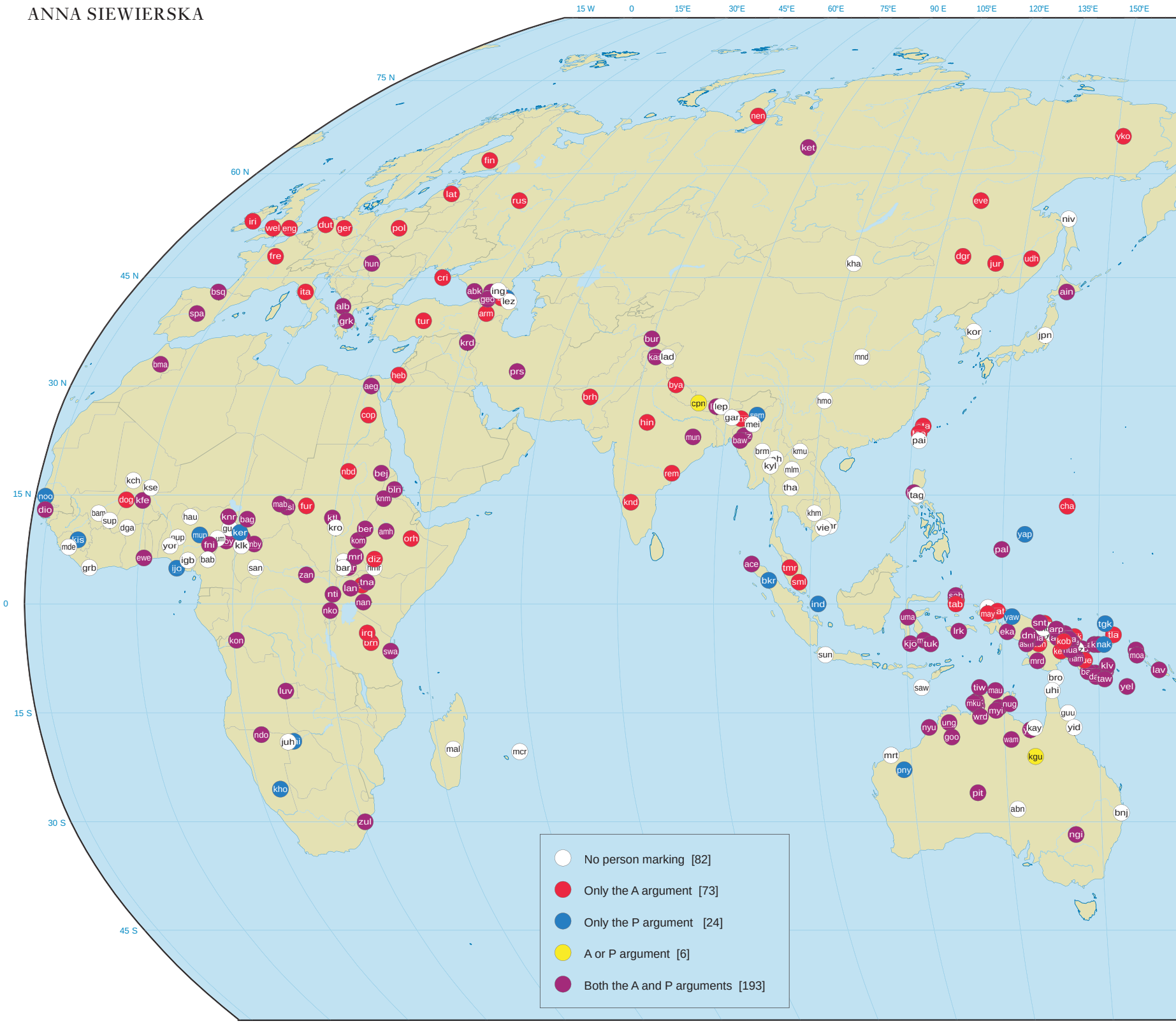
Given that Map 102 is about the marking of person, markers of gender and/or number markers such as those in Archi and Khanty have not been considered.

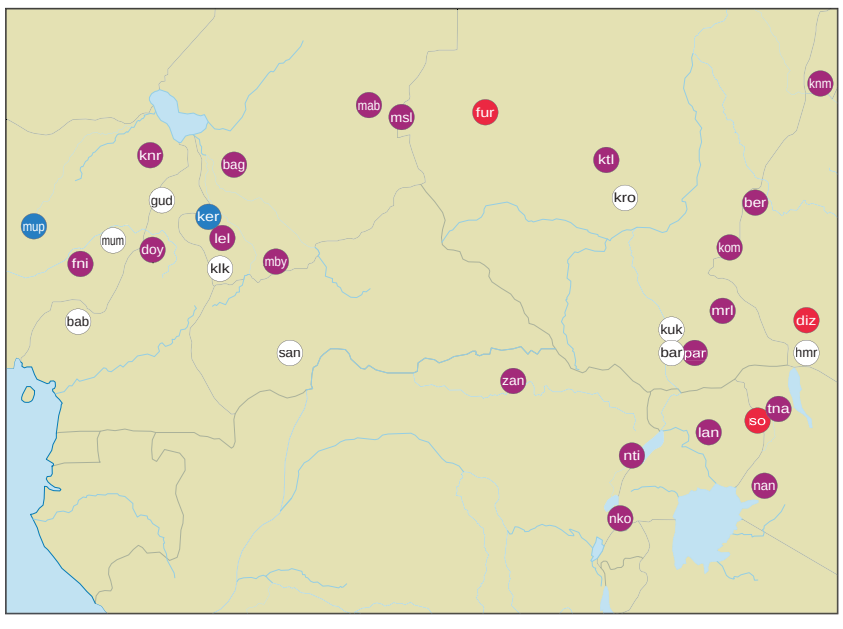
2 Geographical distribution

We see that verbal person marking is very common and widely attested cross-linguistically. In fact, only about a fifth of the languages in the sample lack any form of verbal person marking; such languages are found mainly in West Africa, the Caucasus, and South and South-East Asia. Lack of verbal person marking is particularly rare in North America and New Guinea. Among the languages which manifest verbal person marking, around two thirds exhibit marking of both of the transitive arguments, the A and the P. This is the dominant pattern of person agreement in every area except Eurasia. It is particularly common in North America, New Guinea, and Australia. In languages manifesting person marking of only one of the arguments of a transitive clause, this argument is typically the A (25 per cent) rather than the P (7 per cent). Verbal person marking of just the A is most common in Eurasia. It is characteristic of most of the Indo-European languages of Europe, the Uralic languages of Russia, Finland, Estonia, and Hungary, the Turkic languages, and the Dravidian languages of India. It is also found in eastern Africa, especially among the Afro-Asiatic languages, in the north of South America, and in New Guinea. By contrast, it is extremely rare in North America and there is no instance of it among the languages in the sample from Australia. Verbal person marking of just the P is an infrequent phenomenon everywhere. It is manifested mainly in Africa and South America but not at all in North America, Mesoamerica, or Australia. In Africa it is found mainly among the Chadic languages, in a few Khoisan languages, and sporadically in several groups of Niger-Congo languages.

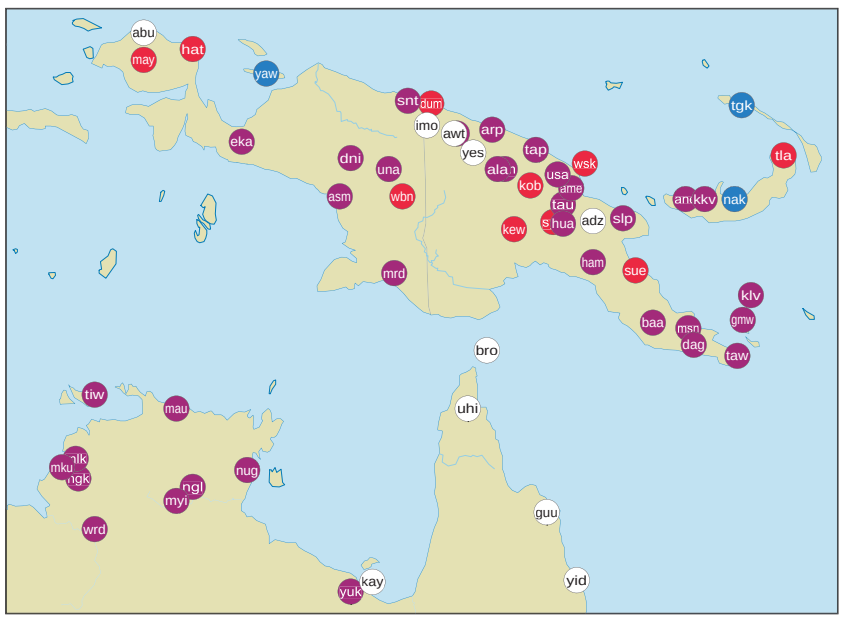
102 Verbal Person Marking

ANNA SIEWIERSKA

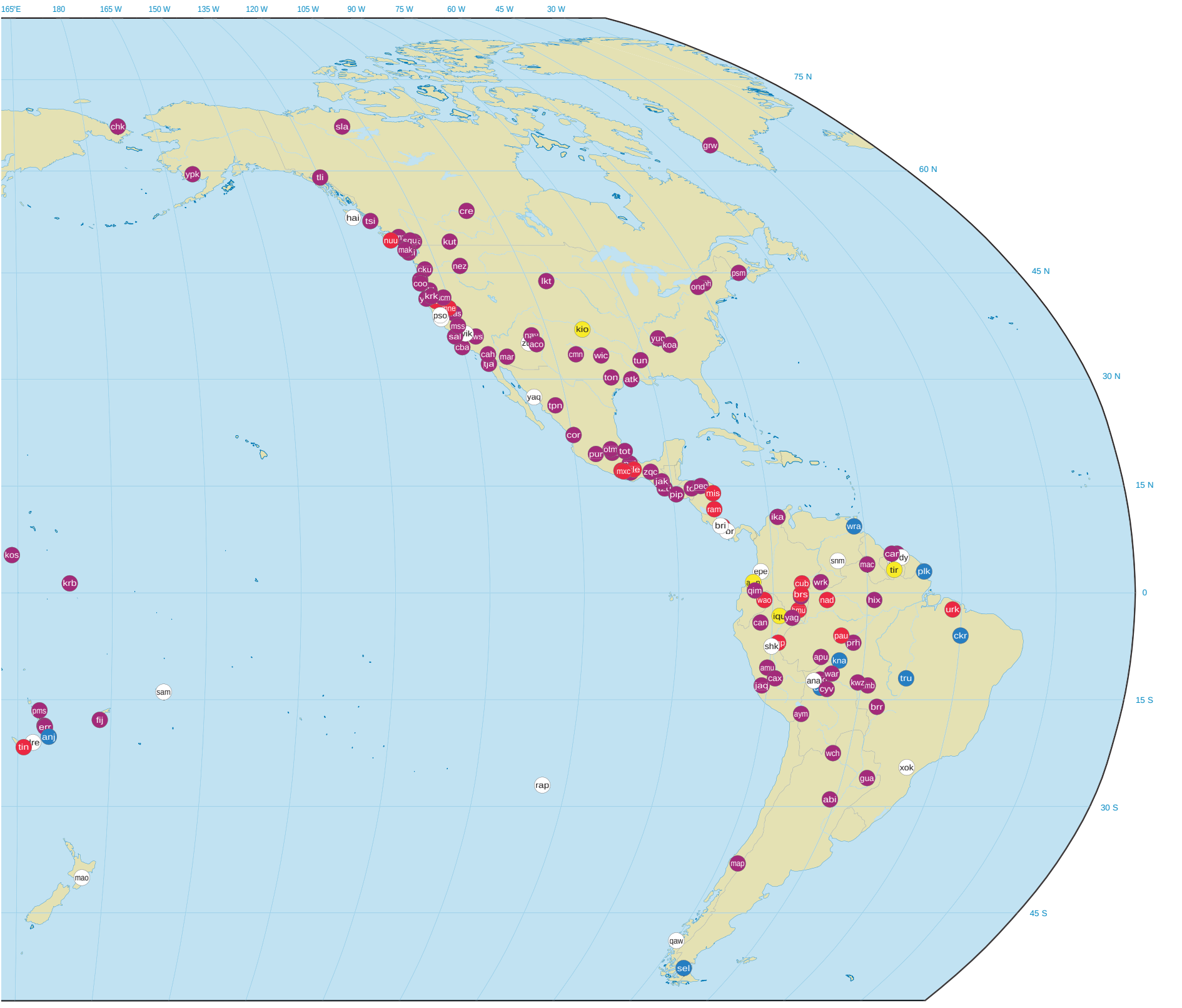


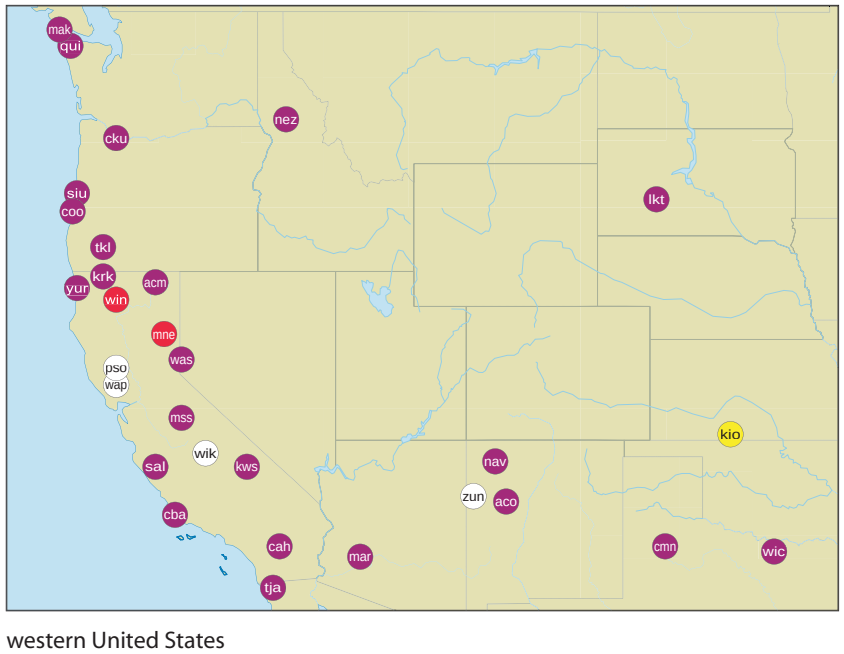


central Africa

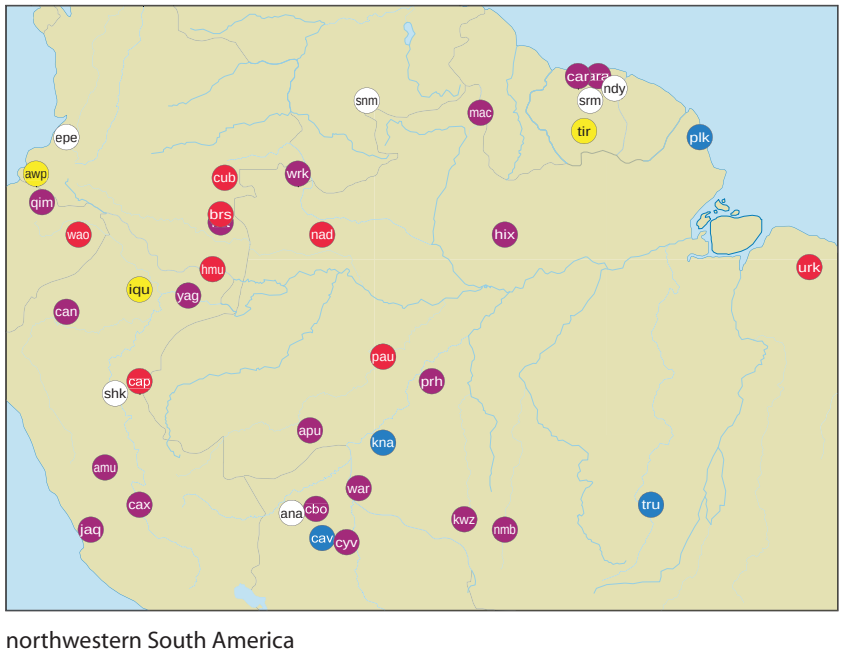


New Guinea and northern Australia





western United States



northwestern South America

103 Third-Person Zero of Verbal Person Marking

ANNA SIEWIERSKA

1 Defining the values

Map 103 represents the distribution of third-person zeroes among verbal person markers of the sole argument of an intransitive clause (i.e. of the S argument). I have chosen to depict the person marking of the S rather than of the transitive A or P, as the verbal person markers of the latter are sometimes fused with each other (see Chapter 104), which makes it difficult to decide which one has a zero realization. The zero realization of third-person S markers is captured by means of six values:

○	1. No person marking of the S	96
○	2. No zero realization of third-person S forms	181
○	3. Zero realization of some third-person singular S forms	21
○	4. Zero realization of all third-person singular S forms	45
○	5. Zero realization of all third-person S forms/No third-person S forms	36
○	6. Zero realization only of third-person nonsingular S	1
	total	380

Most of the languages subsumed under the first value have **no verbal person marking** at all. Some may, however, evince verbal person marking of just the P, as is the case in Palikur (Arawakan; French Guiana) and Yapese (Western Malayo-Polynesian; Caroline Islands) (see Chapter 102, value three).

Given that person markers typically combine person distinctions with those of number and, less often, gender, the size of the inventory of a language's third-person S forms depends on how many number and/or gender distinctions it exhibits in the third person. Thus Kilivila (Western Oceanic; Trobriand Islands, Papua New Guinea), which has a simple singular/plural opposition in the third-person S, possesses just two forms. Tanimbili (Remote Oceanic; Utupua Island) with its singular/dual/plural opposition has three forms (this language is not included in the sample). And Larike (Central Malayo-Polynesian; Ambon Island), which exhibits a singular/dual/trial/plural contrast plus a gender opposition in the singular and plural, has six basic forms:

(1) Kilivila (Senft 1986: 36)

	SG	DU	PL
1INCL		*ta-*	*ta- -s*
1EXCL	*a-*	*ka-*	*ka- -si*
2	*ku-*		*ku- -si*
3	*i-/e-*		*i- - si*

(2) Tanimbili (Tryon 1994: 628)

	SG	DU	PL
1INCL		*si-*	*misu-*
1EXCL	*ŋyi-*	*me-*	*misu-*
2	*nu-*	*mwa-*	*muku-*
3	*i-*	*ŋgi(li)-*	*ŋgu-*

(3) Larike (Laidig 1993: 321)

	SG	DU	TRI	PL
1INCL		*itua-*	*itidu-*	*ite-*
1EXCL	*au-*	*arua-*	*aridu-*	*ami-*
2	*a-/ai-*	*irua-*	*iridu-*	*imi-*
3.HUM	*ma-/mei-*	*matua-*	*matidu-*	*mati-*
3NONHUM	*i-*			*iri-*

Languages may also have separate verbal person paradigms dependent on tense, aspect, mood, and even polarity (e.g. Salinan). Needless to say, this too may result in a greater number of third-person forms. For instance, in Amele (Trans-New Guinea; Papua New Guinea) there are eight classes of person markers, used in different tenses, aspects, and moods. As shown in (4), though in the dual and plural the second and third person are nondistinct, there are still twelve different third-person forms.

(4) Amele (Roberts 1987: 277–8)

	1	2	3	4	5	6	7	8
1SG	*-ig*	*-ig*	*-ig*	*-ig*	*-min*	*-m*	*-em*	*-em*
2SG	*-g*	*-g*	*-g*	*-g*	*-m*	*-m*	*-em*	*-em*
3SG	*-i*	*-Ø*	*-igi*	*-i*	*-b*	*-b*	*-n*	*-Ø*
1DU	*-w*	*-w*	*-w*	*-w*	*-hul*	*-h*	*-h*	*-h*
2/3DU	*-si*	*-si*	*-was*	*-was*	*-bil*	*-b*	*-sin*	*-sin*
1PL	*-q*	*-q*	*-q*	*-q*	*-mun*	*-m*	*-m*	*-m*
2/3PL	*-eig*	*-eig*	*-qag*	*-w*	*-bil*	*-b*	*-ein*	*-ein*

The second value in Map 103 represents languages in which **all the third-person forms that the language distinguishes are overtly realized**, as is the case in Kilivila, Tanimbili, and Larike, though not Amele.

As Amele illustrates, zero realization is primarily associated with the third person singular as opposed to the nonsingular categories. Languages seen as having a zero for the third person singular fall into two subtypes. To the first subtype, represented by value three, belong **languages in which only some realizations of the third person singular are zero while other realizations are not**. Amele is a case in point. Lango (Nilotic, Nilo-Saharan; Uganda) is another. Lango has two paradigms of S (and A) person markers, an "A" set used in the perfective, habitual, and subjunctive and a "B" set used in the progressive. As shown in (5), the third person singular has a zero realization only in the habitual.

(5) Lango (Noonan 1992: 91)

	SG		PL	
	Set A	Set B	Set A	Set B
1	*ǎ-*	*â-*	*ǒ-*	*ô-*
2	*ǐ-*	*î-*	*ǐ-wùnú*	*î-wunu*
3	*ò-*(PERF,SUBJ)	*à-*	*ǐ*	*ǐ-*
	Ø-(HAB)			

To the second subtype, represented by value four, belong **languages in which the third person singular is always zero**, such as Chepang (Kiranti, Tibeto-Burman; Nepal).

(6) Chepang (Caughley 1982: 54–5)

	SG	DU	PL
1INCL		*-ŋə-cə*	*-ŋ-sə*
1EXCL	*-ŋa*	*-təyh-cə*	*-təyh-ʔi*
2	*-naŋ*	*-naŋ-jə*	*-naŋ-sə*
3	*Ø*	*-cə*	*-ʔi/sə*

The fifth value represents **languages which have overt verbal person forms for the first and second person but not for the third person (neither singular nor plural)**, as is the case in So (Kuliak, Nilo-Saharan; Uganda).

(7) So (Carlin 1993: 79)

	SG	PL
1INCL		*-(i)ine*
1EXCL	*-(ɪ)sa*	*-(i)ise*
2	*-(ɪ)ba*	*-(i)ide*
3	*Ø-*	*Ø-*

Another way of describing such languages is to say that they exhibit no third-person forms at all. The description in terms of zero realization is chosen here because we are comparing these languages with others showing values two to six.

The final, sixth value covers **languages which have zero marking of the third person solely in some nonsingular number**. The only instance of such zero marking that I am aware of (apart from languages such as English or Trumai which have overt person marking only in the third person singular) involves gender distinctions as well: this is Barasano (Tucanoan; Brazil and Colombia), where there is a zero marker for third-person plural inanimates.

(8) Barasano (Jones and Jones 1991: 73–4)

1SG	*-ha*	1PL	*-ha*
2SG	*-ha*	2PL	*-ha*
3SG.M	*-bĩ*	3PL.ANIM	*-bã*
3SG.F	*-bõ*	3PL.INAN	*-Ø*
3SG.INAN	*-ha*		

2 Geographical distribution

Map 103 illustrates that overt realization of third-person S markers is the cross-linguistic norm, as nearly two-thirds of the languages in the sample display such marking. What we also see, however, is that there are considerable areal differences with regard to the presence of third-person zeroes. In South-East Asia, the Pacific, and New Guinea, overt third-person forms are clearly favoured over zeroes. This may also be observed in Africa and somewhat less clearly in Eurasia. In Australia and the Americas, on the other hand, there are slightly more languages with some form of zero marking in the third person than languages with only overt third-person forms. Within the Americas, zero marking is more common in North America than in South America, and is particularly frequent in Mesoamerica.

Significant areal differences may also be observed in regard to the type of zero marking displayed. In North America complete lack of third-person markers is encountered more frequently than a zero just in the third singular or just a third-person zero allomorph. In Australia and Eurasia, the converse is the case. And in South America the instances of zero marking are more or less evenly distributed over three types of zero marking, namely absence of third-person forms, zero in third singular, and only a zero allomorph.

3 Theoretical issues

Several lines of explanations have been advanced for the existence of third-person zeroes. Some scholars treat the issue as an instance of loss and/or reanalysis of previously overt markers, others as a case of failure of the third-person markers to develop.

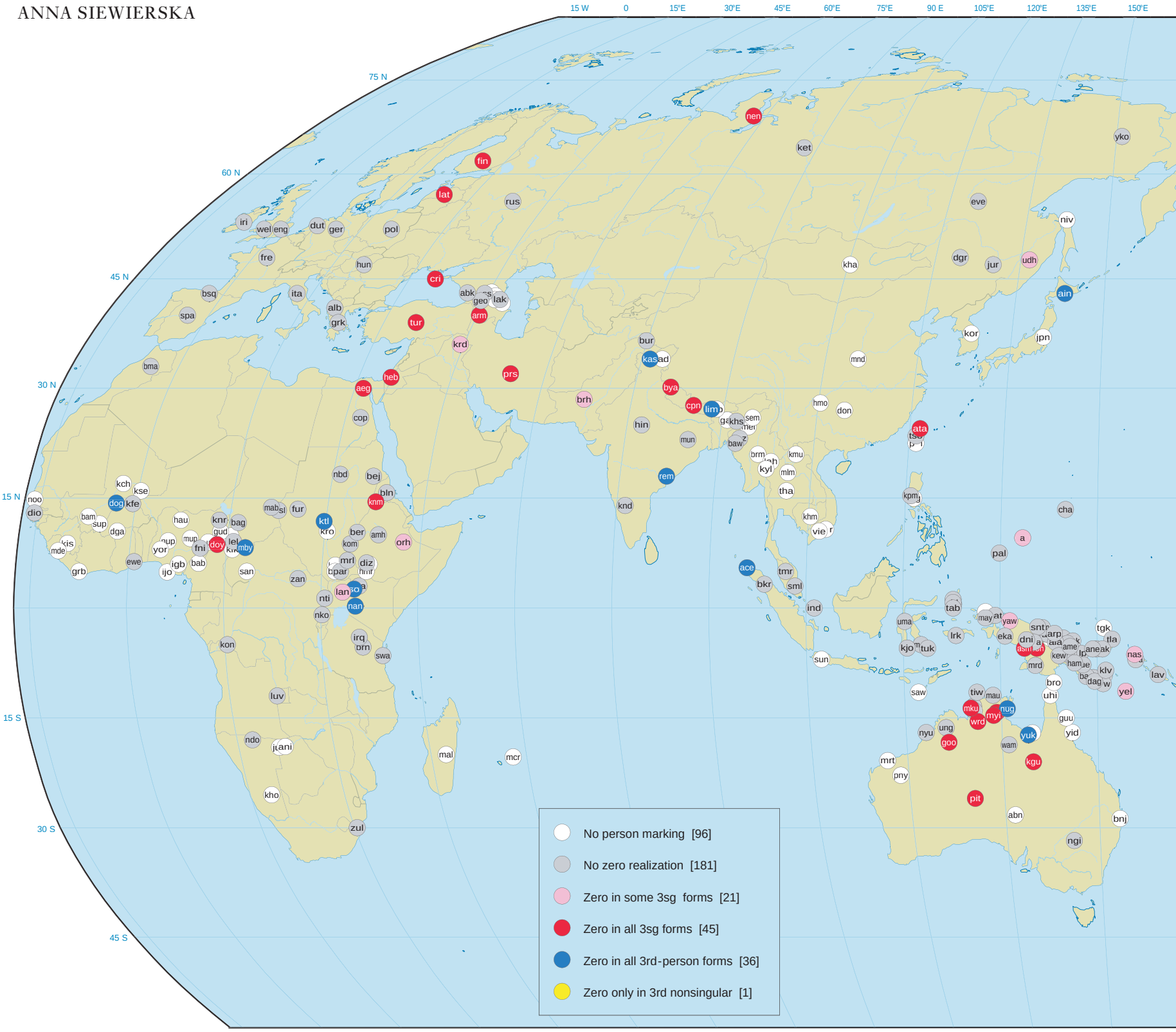
The most widely accepted explanation of the first kind attributes the zero marking of third-person verbal forms to the principle of economy. The principle of economy favours one of the exponents of a paradigm being non-overt. That this should be the third person rather than the second or the first is seen to follow from the higher frequency of first-person verbal forms in discourse than of second- or first-person forms. The effect of frequency on the form of linguistic expressions has long been recognized and is captured in what is commonly referred to as Zipf's law, i.e. "high frequency is the cause of small magnitude" (Zipf 1935: 29). In other words, it is the tendency for speakers to shorten the linguistic expressions used most commonly that motivates the existence of zero realizations of the third person. For some discussion of this position, see Haiman (1985) and Croft (1990).

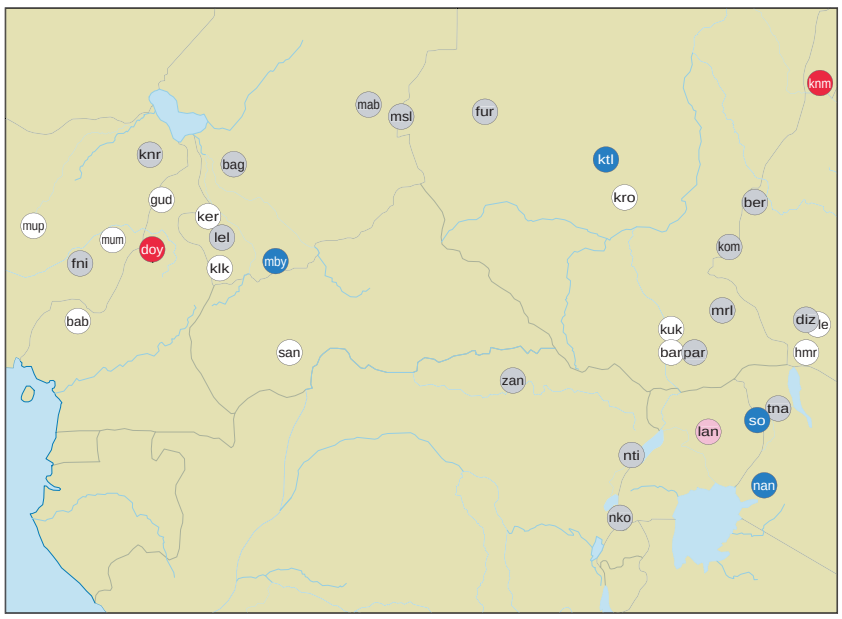
An alternative explanation for third-person zeroes has been developed by Koch (1995), who argues that there is a strong tendency in languages to reinterpret third-person verbal forms as part of the stem or as tense markers. This tendency, he claims, is due to the pressure of iconicity, i.e. the preference for morphological structure to mirror cognitive structure. Adopting the view, most strongly articulated by Benveniste (1971), that the third person is cognitively a "nonperson", and therefore unmarked vis-à-vis the first and second person, he argues that it should therefore also be unmarked morphologically.

In contrast to the above two explanations, Ariel (2000) attributes third-person zeroes to the fact that third-person forms, unlike first- and second-person ones, simply did not develop. In accessibility theory, which constitutes the context of her explanation, the coding of discourse referents is seen to reflect speakers' assumptions as to the degree of accessibility of the referents in the memory store of the addressee. The claim is that the higher the accessibility of the referent, the more attenuated its encoding. Verbal person markers are considered to be high-accessibility coding devices, second only to zero and reflexives. Thus, according to accessibility theory they should be used as the means of referent encoding only in the case of highly accessible referents. According to Ariel, first- and second-person pronouns meet this criterion, but third-person pronouns do not. In other words, the referents of third-person pronouns tend to be insufficiently accessible to warrant phonological reduction, cliticization, and affixation. Consequently, third-person verbal markers, in contrast to first- and second-person ones, tend not to arise. A critical assessment of this view is provided in Siewierska (2004), based in part on the fact that it is primarily the third person singular that is non-overt, rather than the third person in general, as the accessibility explanation would lead one to expect.

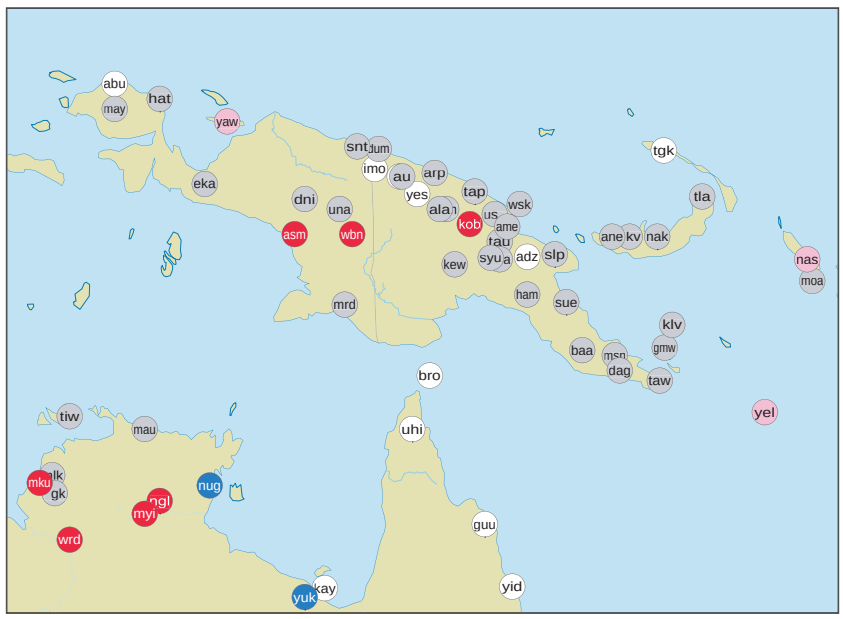
103 Third-Person Zero of Verbal Person Marking

ANNA SIEWIERSKA

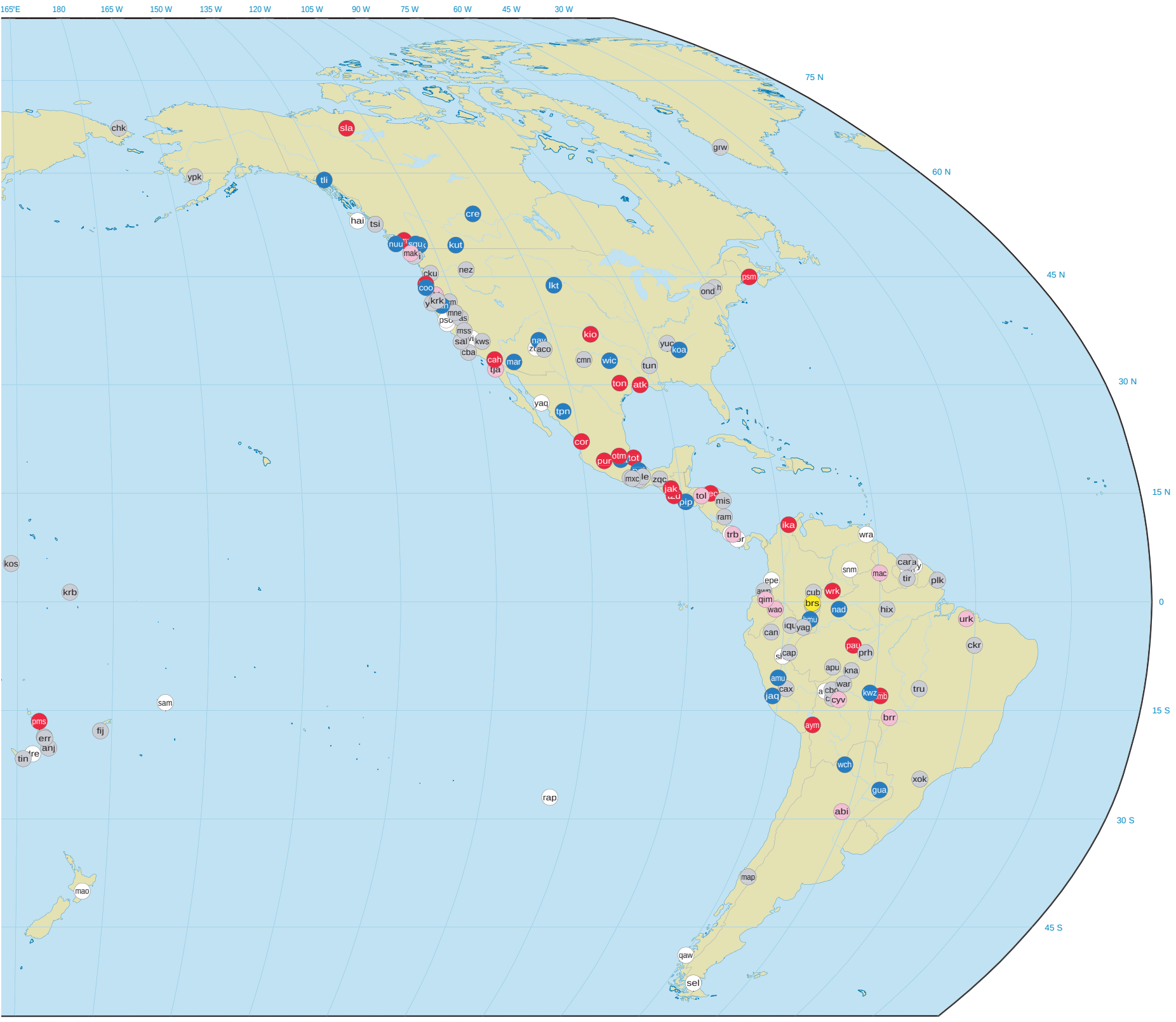


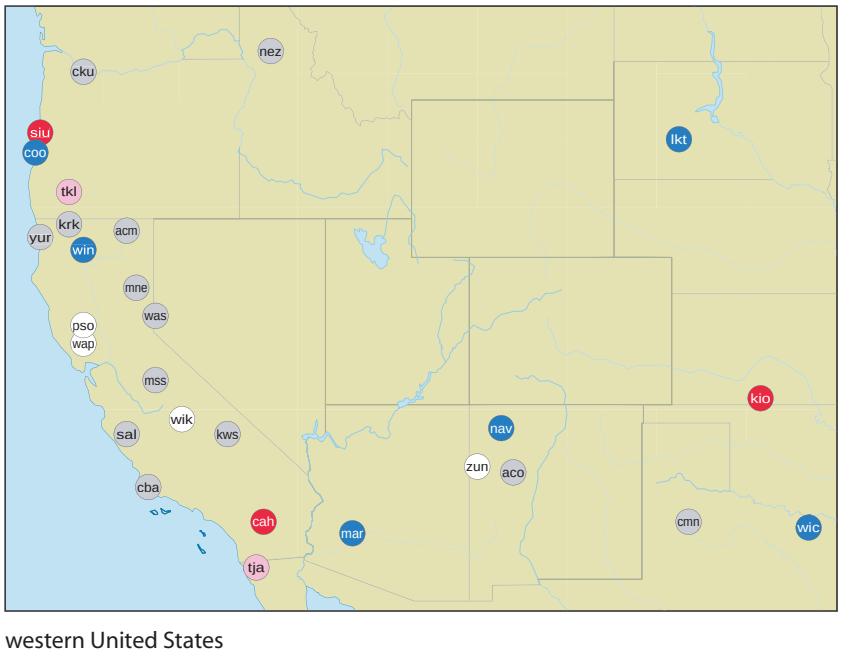


central Africa

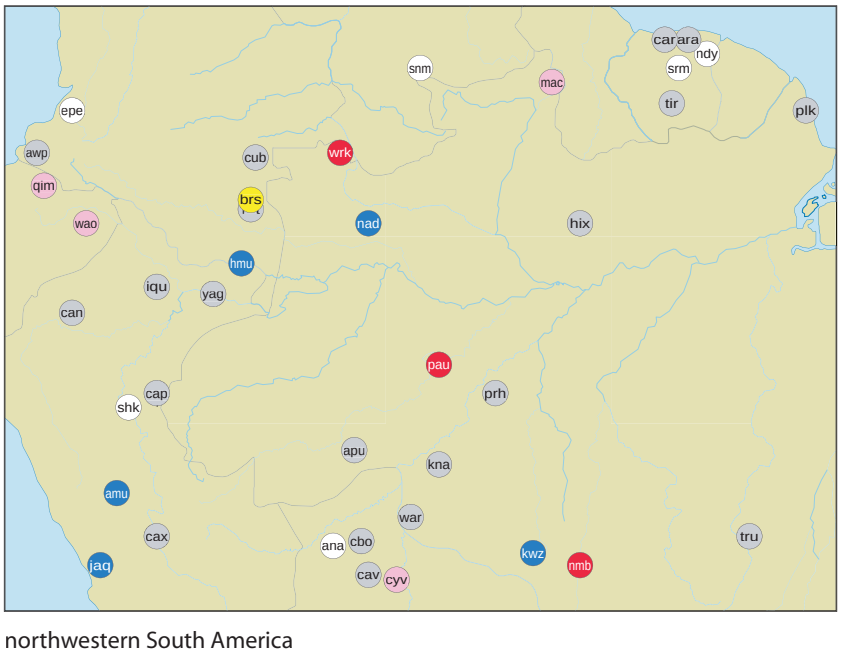


New Guinea and northern Australia





western United States



northwestern South America

104 Order of Person Markers on the Verb

ANNA SIEWIERSKA

1 Defining the values

Map 104 represents the order of A and P person markers on the verb relative to each other, where A stands for the person marker of the agentive argument of a transitive verb and P for the person marker of the patient argument of the transitive verb. Five values are represented:

○	1. A and P do not, or do not both, occur on the verb	187
●	2. A precedes P	96
●	3. P precedes A	57
●	4. Both orders of A and P occur	19
●	5. A and P are fused	20
	total	379

The first value covers languages that exhibit no verbal person marking, have marking of only one of the transitive arguments, or allow both arguments to be marked but not at the same time. This last possibility is sometimes found in languages which display hierarchical alignment of verbal person forms (see Chapter 100), as is the case in Guajajara (Tupi-Guarani; northeastern Brazil). As shown in (1), in Guajajara the transitive verb may agree with either the A as in (1a) or the P as in (1b), depending on which is higher on the hierarchy: 1 > 2 > 3, but it never agrees with both A and P.

(1) Guajajara (Harrison 1986: 419)

a. *a-esak* *kakwez* *kaʔi* *ihe*
1SG.A-see DIST.PST.ATTESTED monkey I
'I saw the monkey.'

b. *he-kisi* *takihe-pupe* *aʔe*
1SG.P-cut knife-with he
'He cut me with a knife.'

The form of the actual agreement markers is sensitive to whether the A outranks the P or vice versa. Thus the form of the first-person marker in (1a), where the A is higher on the hierarchy than the P, is different from that in (1b), where it is the P that is higher than the A.

The second value, **A before P**, subsumes three types of languages: those in which both the A and P person markers are prefixes, as in Swahili (see 2), those in which they are both suffixes, as in Amharic (Semitic; Ethiopia) (see 3), and those in which the two markers occur on opposite sides of the stem, as in Barbareño Chumash (Chumashan; California) (see 4).

(2) Swahili (Ashton 1947: 42)
ni-li-mw-ona
1SG.A-PST-3SG.P-see
'I saw him.'

(3) Amharic (Leslau 1995: 418)
näggär-ä-h
told-3SG.A-2SG.F.P
'He told you.'

(4) Barbareño Chumash (Wash 2001: 43)
kʰ-utiy-in
1SG-see-2SG
'I see you.'

Analogously, the reverse order, **P before A**, covers the same three types of markers: A and P prefixes, as in Retuarã (Tucanoan; Colombia) (see 5), A and P suffixes, as in Lillooet (Salishan; British Columbia) (see 6), and a P prefix and A suffix, as in Tauya (Madang, Trans-New Guinea; Papua New Guinea) (see 7).

(5) Retuarã (Strom 1992: 219)
sa-ki-baʔa-koʔo
3SG.P-3SG.A-eat-PST
'He ate it.'

(6) Lillooet (van Eijk 1997: 45)
núkʼʷʔ-an-c-as
help-TR-1SG.P-3SG.A
'He helped me.'

(7) Tauya (MacDonald 1990: 118)
nen-yau-a-ʔa
3PL.P-see-3SG.A-IND
'She/he saw them.'

In the vast majority of languages which have been assigned value two or three, the A and P person markers exhibit a unique location relative to the verbal stem: both are prefixes, or both are suffixes, or one is a prefix and the other a suffix. There are, however, languages in which the A and P exhibit the same order relative to each other, but vary with respect to their location vis-à-vis the stem. For instance, in Coeur d'Alene (Salishan; northern Idaho), one set of A and P forms is used in the completive and stative aspects, and another in the progressive. The first set of forms are suffixes, the second are prefixes. As shown in (8), the order of the A and P forms remains the same: the P forms precede the A forms.

(8) Coeur d'Alene (Kroeber 1986: 77)

a. *gʷíč-t-s-ən*
see-TR-2SG.P-1SG.A
'I saw thee.'

b. *kʷ-i-ʔc-gʷíč-əm*
2SG.P-1SG.A-ASP-see-MID
'I am seeing thee.'

Such languages have been classifed under values two or three, depending on whether the order is A > P or P > A, respectively.

Turning to the fourth value, **both orders of the A and P**, an example of a language displaying both orders of the A and P relative to each other is Ika (Chibchan; northeastern Colombia). In Ika the P forms and the second-person singular and first- and second-person plural A forms are prefixes while the first singular A marker in the past and irrealis is a suffix. The third-person forms are zero. In the case of prefixal forms the A precedes the P (9a), but when the A is first person, the order is P before A (9b).

(9) Ika (Frank 1990: 52, 21)

a. *nA-niwe-ʔzasana* *ki* *u-š-e*
2SG.A-1PL.P-pay CNTR AUX-MED-Q
'Did you pay us?'

b. *mi-tšua-na-rua*
2SG.P-see-DIST-1SG.A
'I saw you.'

In some languages the alternative orders of the A and P relative to each other are dependent on the position of the referents of the markers on the person and/or animacy hierarchies. This is not uncommon among the languages of Australia (e.g. Bininj Gun-Wok, Yukulta, Yulparija). Such hierarchically based ordering of the A and P is also found in Yimas (Lower Sepik-Ramu; Papua New Guinea), in which two hierarchies are actually involved, a person hierarchy (1 > 2 > 3) and a role hierarchy. The latter ranks the P higher than the A in the case of the first and second person, and the A higher than the P in the case of the third person. The higher ranking participant is placed closer to the verb stem. Thus when both

participants are third person or the A is first or second person and the P third, we have PA order as in (10).

(10) Yimas (Foley 1991: 202, 205)
a. *pu-n-tay*
3PL.P-3SG.A-see
'He saw them.'
b. *pu-ka-tay*
3PL.P-1SG.A-see
'I saw them.'

But when the A is third person and the P first or second, or the A is second person and the P first, we have AP order, as in (11).

(11) Yimas (Foley 1991: 205–6)
a. *pu-ŋa-tay*
3PL.A-1SG.P-see
'They saw me.'
b. *ma-ŋa-tay*
2SG.A-1SG.P-see
'You saw me.'

The conflict between the two hierarchies which arises when there is a first-person A and a second-person P is resolved by means of a portmanteau morpheme *mpan-/kampan*, as illustrated in (12).

(12) Yimas (Foley 1991: 207)
kampan-tay
1SG.A.2SG.P-see
'I saw you.'

In languages in which both the A and P forms occur on the same side of the verb, it is not always possible to determine which part signals the A and which part the P because the two are **fused**. Consider, for instance, the verbal paradigm in (13), involving the verb */ill/* 'to see' in the present tense in Jaqaru (Aymaran; Peru).

(13) Jaqaru (Hardman 2000: 57)

ill-k-ima	I see you	1SG > 2SG
ill-k-uta	you see me	2SG > 1SG
ill-k-ushta	you see us	2SG > 1PL
ill-k-utu	she sees me	3SG > 1SG
ill-k-ushtu	she sees us	3SG > 1PL
ill-k-tma	she sees you	3SG > 2SG
ill-k-ta	you see him	2SG > 3SG
ill-k-t"a	I see him	1SG > 3SG
ill-k-tna	we see him	1PL > 3SG
ill-k-ti	she sees him	3SG > 3SG

A closer look at the above forms reveals that the second person is always associated with *-ma* if it is a P and with *-ta* if it is an A. But beyond this, no further definite segmentation appears to be possible.

2 Geographical distribution

As indicated on the map, among the 192 languages which have verbal person markers for both the A and P, there is a preference for placing the A before the P. This preference, however, is not manifested in all areas of the world. The placement of A before P is particularly frequent in South-East Asia and the Pacific. It is also very common in Africa. Interestingly, all the African languages in the sample that display the opposite order (i.e. PA) are Nilo-Saharan. AP order also prevails over PA in Australia, South America, and to a lesser extent in Eurasia.

In contrast to South America, North America exhibits a preference for PA rather then AP order. AP order is not only disfavoured relative to PA but also more geographically restricted. We see that PA languages are distributed all over the continent, while most of the AP languages are in the western and central parts of the continent. In Mesoamerica, on the other hand, PA order is just as common as AP. This is also the case in New Guinea.

Fused A and P verbal person markers are found primarily in the Americas, particularly in North America. They are also characteristic of the languages of the far north of Australia. Fusion of A and P markers is uncommon in Africa and Eurasia.

Languages which have both orders of the A and P markers relative to each other are common only in Australia. In fact, in Australia such languages prevail over those displaying a single order, be it AP, fused, or PA. The only other area in which both orders are encountered with some regularity is North America.

The order of the A and P relative to each other is not independent of whether the two are prefixes, suffixes, or occur on opposite sides of the verb. AP order is particularly favoured when the two person markers are on opposite sides of the verb; of the 84 languages which have the two person markers on opposite sides of the verb and display a unique order of the A and P, in 75 per cent the A precedes the P. This tends to be the case in SVO and some verb-initial languages. This explains the frequency of AP order in South-East Asia, the Pacific, and Africa. Among the 36 languages in which the A and P are both prefixes, AP and PA orders are very evenly distributed: 19 languages have AP order and 17 PA. This is also more or less the case in the 34 languages in which the A and P are both suffixes: 15 have AP order and 19 have PA order. Fused markers occur somewhat more frequently in prefixal position (twelve languages) than in suffixal position (eight languages). This also holds for the languages in the sample manifesting alternative orderings of the A and P; thirteen (65 per cent) are prefixing.

Perhaps the most common pattern overall is A-V-P, in keeping with the above discussion.

3 Theoretical issues

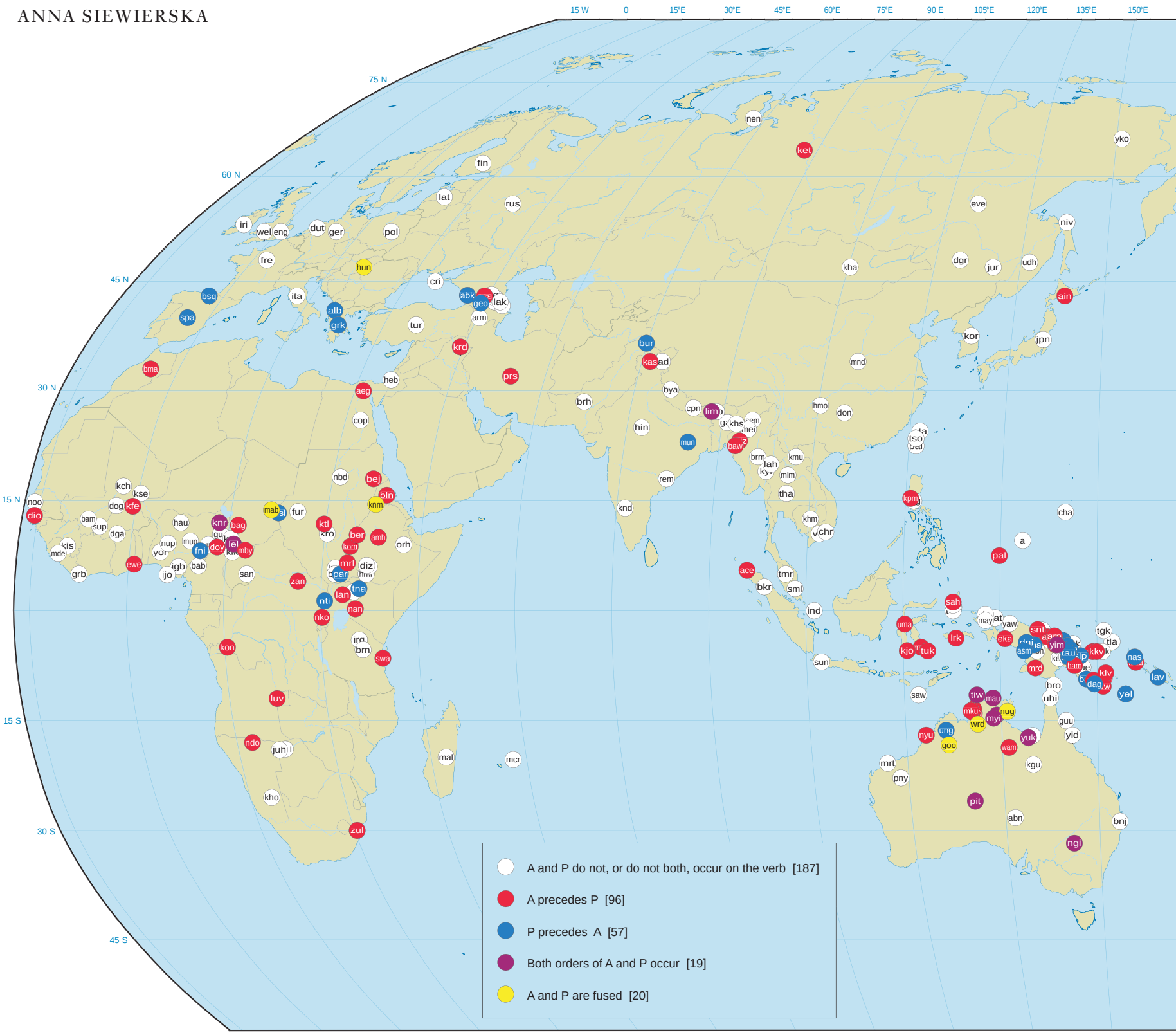
One of the explanations that have been advanced for the ordering of person affixes relative to each other is the principle of relevance (cf. Bybee 1985, Bybee et al. 1990), which defines a preference for placing affixes with a greater semantic effect on the head closer to the head than those exerting a lesser effect. The effect of the principle of relevance on the ordering of the A and P is taken to be dependent on the alignment of verbal person forms (cf. Bittner and Hale 1996). The expectation is that in accusative alignment the P should be positioned closer to the verbal stem than the A, and vice versa in ergative alignment. The predicted ordering patterns are thus the ones shown in (14).

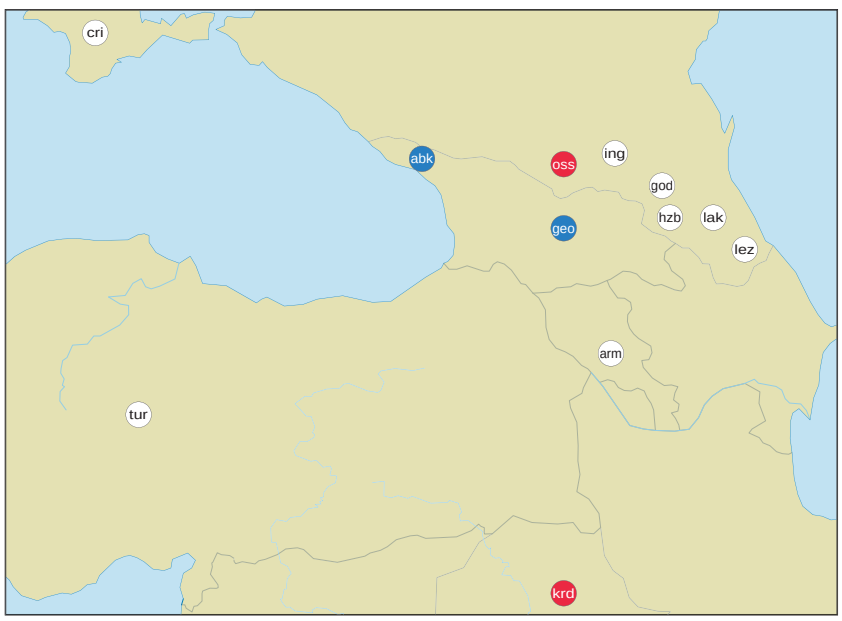
(14) a. NOMINATIVE/ACCUSATIVE A-P-V-P-A
b. ABSOLUTIVE/ERGATIVE P-A-V-A-P

Among the seventy languages in which the A and P are either both prefixes or both suffixes and which have a unique ordering of the two relative to each other, only fifty display either accusative or ergative alignment as the sole alignment. The alignment is accusative in forty-three languages and ergative in seven. The P is placed closer to the verbal stem than the A in 63 per cent of the languages with accusative verbal alignment, and the A is placed closer to the stem than the P in 71 per cent of the languages with ergative alignment. In all, the predicted ordering occurs in 64 per cent of the relevant languages. Another factor which is seen to influence affix order is the degree of grammaticalization of the morphemes in question. Diachronically older forms, i.e. forms that have undergone more development, are expected to occur closer to the stem than younger forms (see e.g. Bybee et al. 1991: 33). Due to lack of diachronic data, the effects of this factor are much more difficult to test.

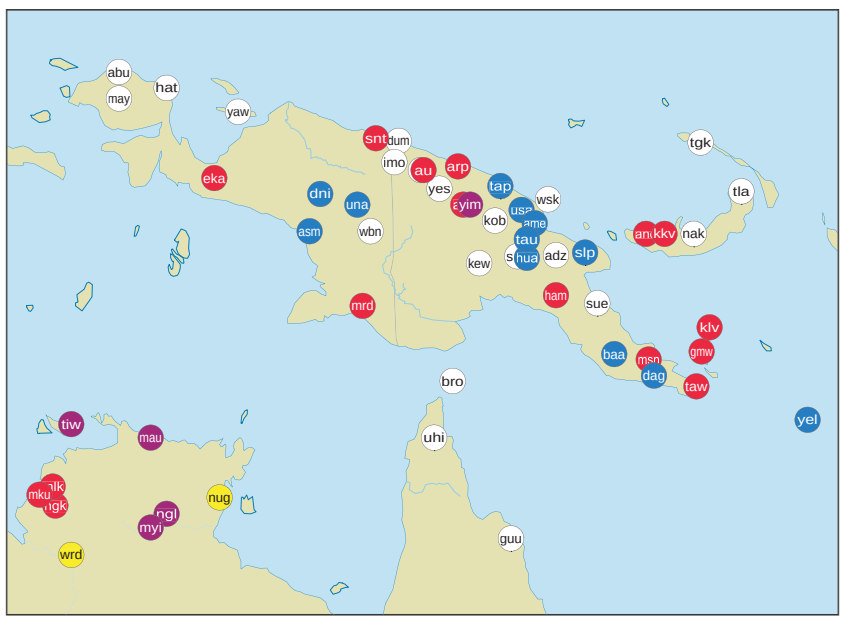
104 Order of Person Markers on the Verb

ANNA SIEWIERSKA

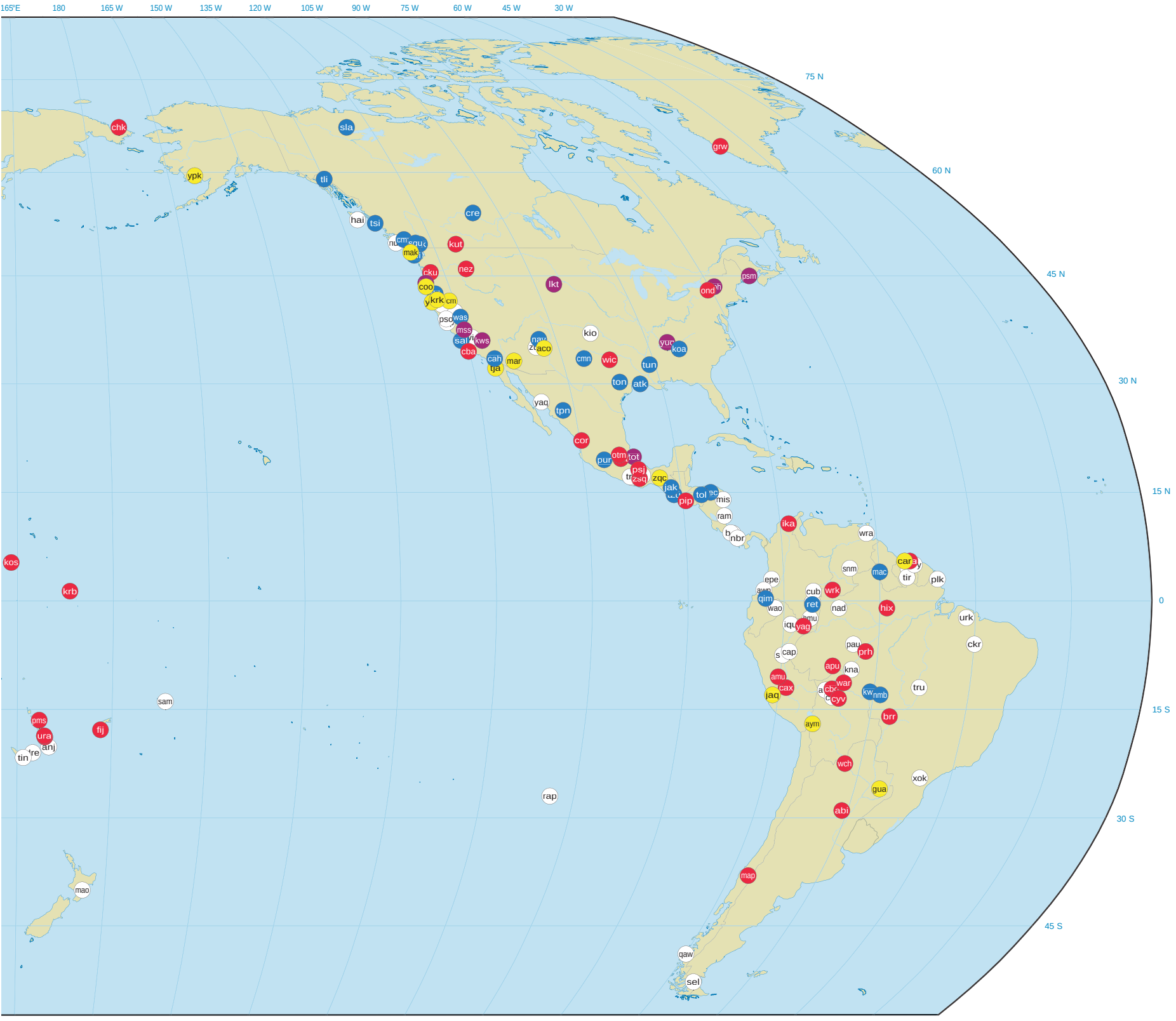


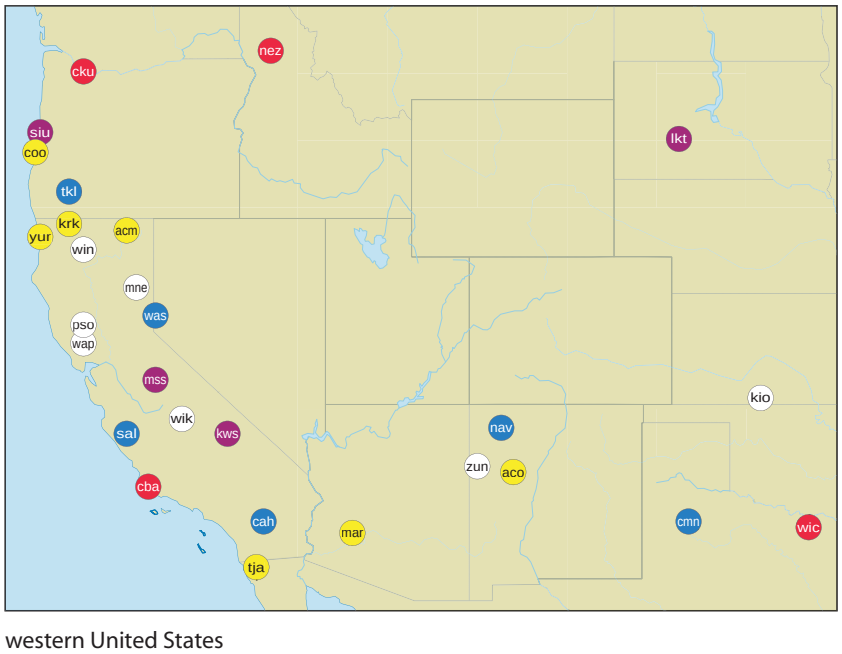


Caucasus and Black Sea area

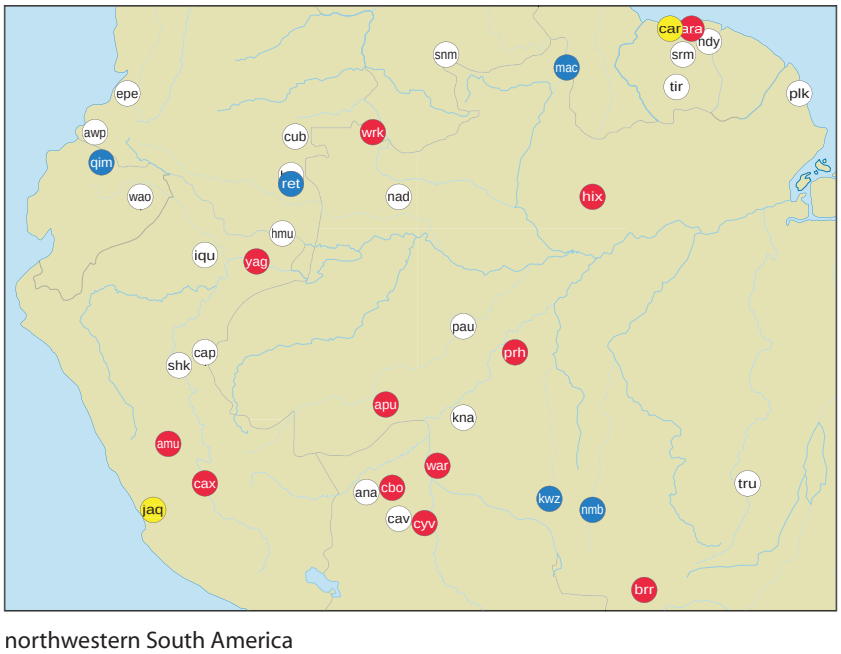


New Guinea and northern Australia





western United States



northwestern South America

105 Ditransitive Constructions: The Verb 'Give'

MARTIN HASPELMATH

1 Overview of the values

Ditransitive verbs are verbs with two arguments in addition to the subject: a "recipient" or "addressee" argument, and a "theme" argument. Typical ditransitive verb meanings are 'give', 'sell', 'bring', 'tell'. Since different ditransitive verbs occur in different constructions in many languages, only a single verb meaning was taken into account for this chapter: the verb 'give', which is probably the most frequent ditransitive verb in all languages.

The parameter of variation examined here is the relationship between the coding properties of the two ditransitive objects and the properties of the single object in a standard monotransitive construction (with a typical transitive verb like 'catch' or 'kill'; this object is called *patient* here). The two coding properties which are taken into account are **case or adpositional marking** and **person-number indexing** ("agreement") on the verb.

There are three main types of ditransitive constructions:

(i) In the **indirect-object construction**, the theme of the ditransitive verb (i.e. the argument expressing the gift) is coded like the monotransitive patient, and the recipient is coded differently. For example, in Krongo (Kadugli; Sudan), the patient lacks case marking (cf. 1a), and so does the ditransitive theme *bìità* 'water' in (1b), whereas the recipient *à-káaw* 'to the person' is marked by a dative-case prefix.

(1) Krongo (Reh 1985: 267–8)

a. *N-àpá-ŋ* *àʔàŋ* *káaw* *ɣ-íkkì.*
1-PFV.hit-TR I person M-that
'I hit that man.'

b. *N-àdá-ŋ* *àʔàŋ* *bìità* *à-káaw.*
1-PFV.give-TR I water DAT-person
'I gave water to the man/woman.'

In these constructions, the monotransitive patient and the ditransitive theme are grouped together as **direct object**, as opposed to the recipient, which is referred to as **indirect object**.

(ii) In the **double-object construction**, both the theme and the recipient of the ditransitive verb are coded like the monotransitive patient. For example, in Panyjima (Ngayarda, Pama-Nyungan; Western Australia) both ditransitive objects bear accusative case, like the monotransitive patient.

(2) Panyjima (Dench 1991: 193)

a. *Ngunha* *parnka* *ngarna-rta* *mantu-yu.*
that lizard eat-FUT meat-ACC
'That lizard will eat the meat.'

b. *Ngatha* *yukurru-ku* *mantu-yu* *yinya-nha.*
I.NOM dog-ACC meat-ACC give-PST
'I gave the dog meat.'

Note that word order is ignored in this chapter, so that the two objects in (2b) are considered as having identical coding properties.

(iii) In the **secondary-object construction**, it is the recipient of the ditransitive verb that is coded like the monotransitive patient, whereas the ditransitive theme is coded differently. For example, in Chamorro (Western Malayo-Polynesian; Guam), a definite monotransitive patient is preceded by an absolutive marker (cf. 3a), and so is the ditransitive recipient 'child' in (3b), whereas the theme is preceded by an oblique marker.

(3) Chamorro (Topping 1973: 241, 251)

a. *Ha* *tuge'* *i* *kannastra.*
he.ERG weave ABS basket
'He wove the basket.'

b. *Ha* *na'i* *i* *patgon* *ni* *leche.*
he.ERG give ABS child OBL milk
'He gave the milk to the child.'

In this construction, the monotransitive patient and the ditransitive recipient are grouped together as **primary object**, as opposed to the theme, which is referred to as **secondary object**.

In addition to these three major construction types, we also need to distinguish a fourth type: languages that show a **mixture of constructions**. For example, in English the verb 'give' can occur both in an indirect-object construction and in a double-object construction (see the translations of (1b) and (2b)), so English is classified as belonging to the mixed type.

Thus, the following four values are shown on the map:

1. Indirect-object construction	189
2. Double-object construction	84
3. Secondary-object construction	66
4. Mixed	39
total	378

There are of course further logical possibilities. In particular, one can imagine a language in which neither the recipient nor the theme behaves like the monotransitive patient. Such further types are unattested.

It should be noted that pronominal objects often behave differently from full-NP objects. For example, in Pero (West Chadic; Nigeria), full-NP recipients are marked with the preposition *ti* 'to' and thus occur in an indirect-object construction (cf. 4a), whereas pronominal recipients are coded in the same way as pronominal themes and patients, so that they occur in a double-object construction (cf. 4b).

(4) Pero (Frajzyngier 1989: 167, 109)

a. *Músà* *mún-kò* *júrà* *tí* *Ɗíllà.*
Musa give-COMP peanut to Dilla
'Musa gave peanuts to Dilla.'

b. *À-mún-tée-nô-té-m.*
NEG-give-VENT-1SG.OBJ-3SG.OBJ-NEG
'He did not give it to me.'

In English, the opposite situation obtains in many varieties: pronominal objects allow only the indirect-object construction (*Give it to them*/**Give them it*). To simplify the picture, only constructions with two full-NP object arguments have been taken into account for this chapter. Thus, Pero has been classified as showing the indirect-object construction, and English as mixed.

2 The indirect-object construction

In almost all indirect-object constructions, the recipient has a specific case marker (as in 1b) or adposition (as in 4a) that distinguishes it from the theme, which may be unmarked or bear a direct-object marker. Constructions in which the recipient is unmarked, contrasting with direct-object marking on the theme, are unattested. There are a very few languages in which the coding difference of the recipient concerns verb indexing, not case or adpositional marking. One such language is Abkhaz (North-West Caucasian; Georgia), where full-NP arguments are not marked but are indexed in the verb by prefixes. The recipient NP is indexed by a special set of recipient prefixes which follow the patient/theme prefixes:

(5) Abkhaz (Hewitt 1979: 209, 105)

a. *A-šʷqʷ'-kʷà* *Ø-z-be-yt'.*
ART-book-PL 3PL.PAT-1SG.AGT-see-FIN
'I saw the books.'

b. *Sarà* *a-xʷəč'-kʷa* *a-šʷqʷ'-kʷà*
I ART-child-PL ART-book-PL
Ø-rə-s-to-yt'.
3PL.THM-3PL.RECP-1SG.AGT-give-FIN
'I gave the books to the children.'

Languages without case or adpositional marking in which only the patient and theme, but not the recipient, are indexed are unattested.

Quite a few languages show **differential object marking**, i.e. a special case marker or adposition on the monotransitive patient (and ditransitive theme) only with animate and/or definite NPs (cf. Bossong 1985, 1998; Comrie 1989: ch. 6). When, as is commonly the case, the differential object marker is identical to the recipient marker, we have a potential problem. Let us look at Spanish, which is well known for its differential object marker *a* occurring on animate patients (cf. 6a–b). The same preposition also marks recipients (cf. 6c).

(6) Spanish

a. *Veo* *el* *libro.*
see.1SG the book
'I see the book.'

b. *Veo* *a* *Juan.*
see.1SG DAT Juan
'I see Juan.'

c. *Le* *doy* *el* *libro* *a* *Juan.*
3SG.OBJ give.1SG the book DAT Juan
'I give the book to Juan.'

Upon surface inspection of these examples, it might appear that both the theme 'book' and the recipient 'Juan' in (6c) behave like the monotransitive patient, so that we would be dealing with a double-object construction. However, examination of a greater range of data shows that with ditransitives the recipient is always marked with *a*, independently of its animacy, while the theme has the same differential marking as the monotransitive patient. Thus, Spanish and similar languages are classified here as indirect-object languages.

3 The double-object construction

In the great majority of cases, the two objects that behave alike and like the monotransitive patient show no marking at all, i.e. neither case marking nor indexing on the verb. Thus, example (7) from Dagbani (Gur, Niger-Congo; Ghana and Togo) is more typical than example (2) above.

(7) Dagbani (Olawsky 1999: 45)
Doo *maa* *ti* *paɣa* *maa* *sima.*
man DEF give woman DEF groundnut
'The man gave groundnuts to the woman.'

The double-object type also includes languages in which both the recipient and the theme are indexed on the verb, and they are indexed by the same markers as the patient. An example is Kinyarwanda (Bantu; Rwanda; Dryer 1983).

In many double-object languages, the recipient and the theme can be distinguished by word order, and when the order of recipient and theme is fixed, it is generally the recipient that comes first (cf. Sedlak 1975, Primus 1998). However, there are also quite a few double-object languages where not even word order distinguishes recipient and theme, so that the role assignment must be inferred from the context (e.g. Shipibo-Konibo (Panoan; Peru), Valenzuela 2002).

4 The secondary-object construction

While there are a few languages that, like Chamorro, show special case marking of the ditransitive theme, in most secondary-object languages the theme's special behaviour concerns verb indexing: In these languages there is no object case marking, and only the monotransitive patient and the ditransitive recipient are indexed on the verb. A typical example comes from Motuna (East Bougainville; Papua New Guinea).

(8) Motuna (Onishi 2000: 117–18)

a. *Aanih* *nii* *tangu-m-u-i-na.*
Aanih (F) me slap-1OBJ-3AGT-NEAR.PST-F
'Aanih slapped me.'

b. *Nii* *ong* *miika* *o-m-i-ng.*
me that betel give-1OBJ-2AGT-PL.IMPF
'Give that betel mixture to me.'

Languages with the opposite pattern, verb indexing of only the ditransitive theme, are unattested.

5 Mixed languages

A language has been classified as mixed if it exhibits two equally widespread alternative constructions with the verb 'give', or if it has two verbs meaning 'give' which occur in different constructions. The former case is illustrated by Mandarin (cf. 9a–b), as well as by English.

(9) Mandarin (Huang and Ahrens 1999: 2)

a. *Wǒ* *sòng* *tā* *yī* *běn* *shū.*
I give s/he one CLF book
'I gave him/her a book.'

b. *Wǒ* *sòng* *yī* *běn* *shū* *geǐ* *tā.*
I give one CLF book give him/her
'I gave a book to him/her.'

There is one further kind of behaviour that could be regarded as mixed: when a construction is characterized both by case marking and by verb indexing, and when these two do not go together. For example, in Bawm (Chin, Tibeto-Burman; Bangladesh, etc.), the recipient is marked by the postposition *sinah*, contrasting with the unmarked theme, and is indexed on the verb by an object clitic which also indexes the monotransitive patient (*nan=* in (10a–b)).

(10) Bawm (Reichle 1981: 149, 152)

a. *Nangmah* *ka* *nan=umpî.*
you I 2=marry
'I marry you.'

b. *Nan* *sinah* *chabu* *ka* *nan=pêk.*
you to book I 2=give
'I give the book to you.'

It has been decided somewhat arbitrarily not to categorize such languages as mixed, but to classify them according to the behaviour of their case or adpositional marking. Thus, Bawm is classified as an indirect-object language for the purposes of this map.

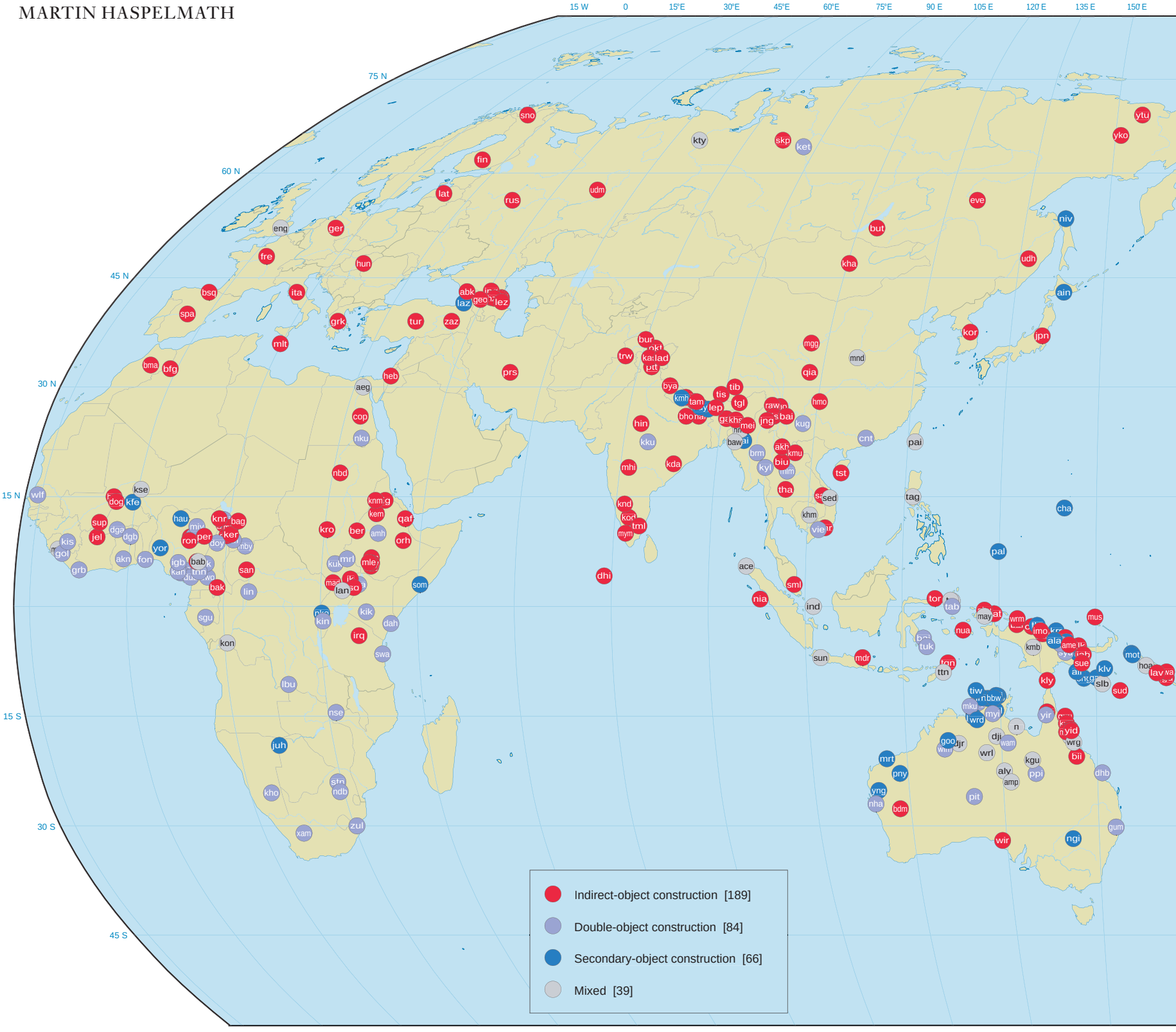
6 Geographical distribution

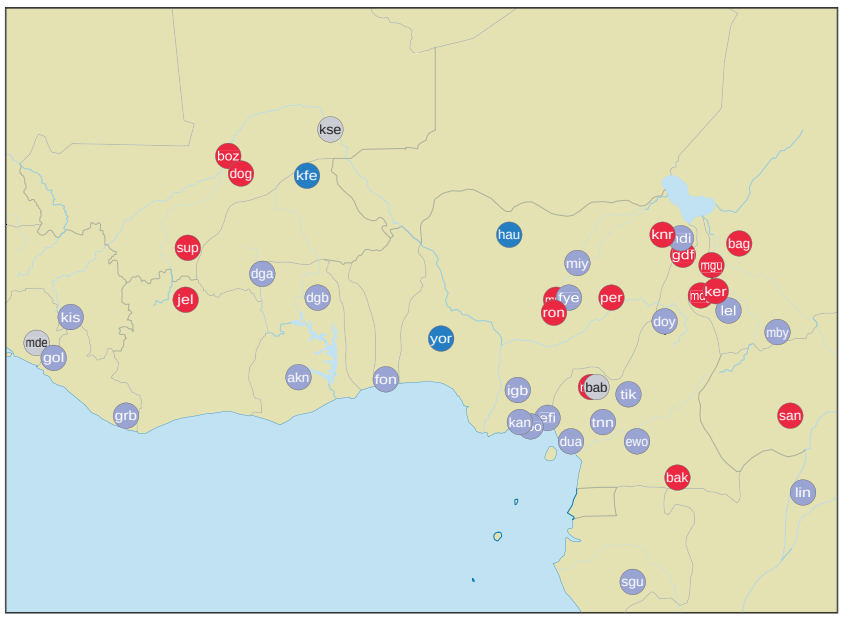
The two clear geographical generalizations are the overwhelming predominance of the indirect-object pattern in Eurasia (except for mainland South-East Asia), and the absence of this pattern from North America. Mesoamerica and adjacent areas in South America are also dominated by indirect-object languages, but elsewhere the patterns are more mixed. In Africa, there appears to be some alignment with language families: Afro-Asiatic and Mande languages tend to show the indirect-object construction, while Niger-Congo languages other than Mande tend to show double-object and secondary-object constructions.

Ditransitive constructions are not independent of other typological parameters: dependent-marking languages with case marking tend to show the indirect-object pattern, head-marking languages with verb indexing of two core arguments tend to show the secondary-object pattern, and the double-object pattern is most common in languages with no case marking (see the maps in Chapters 23, 51, and 104).

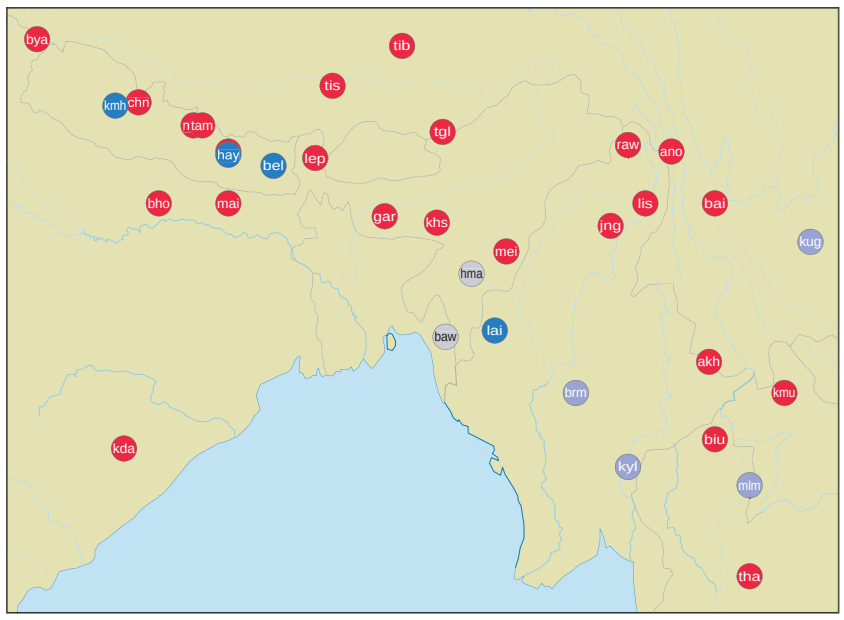
105 Ditransitive Constructions: The Verb ‘Give’

MARTIN HASPELMATH

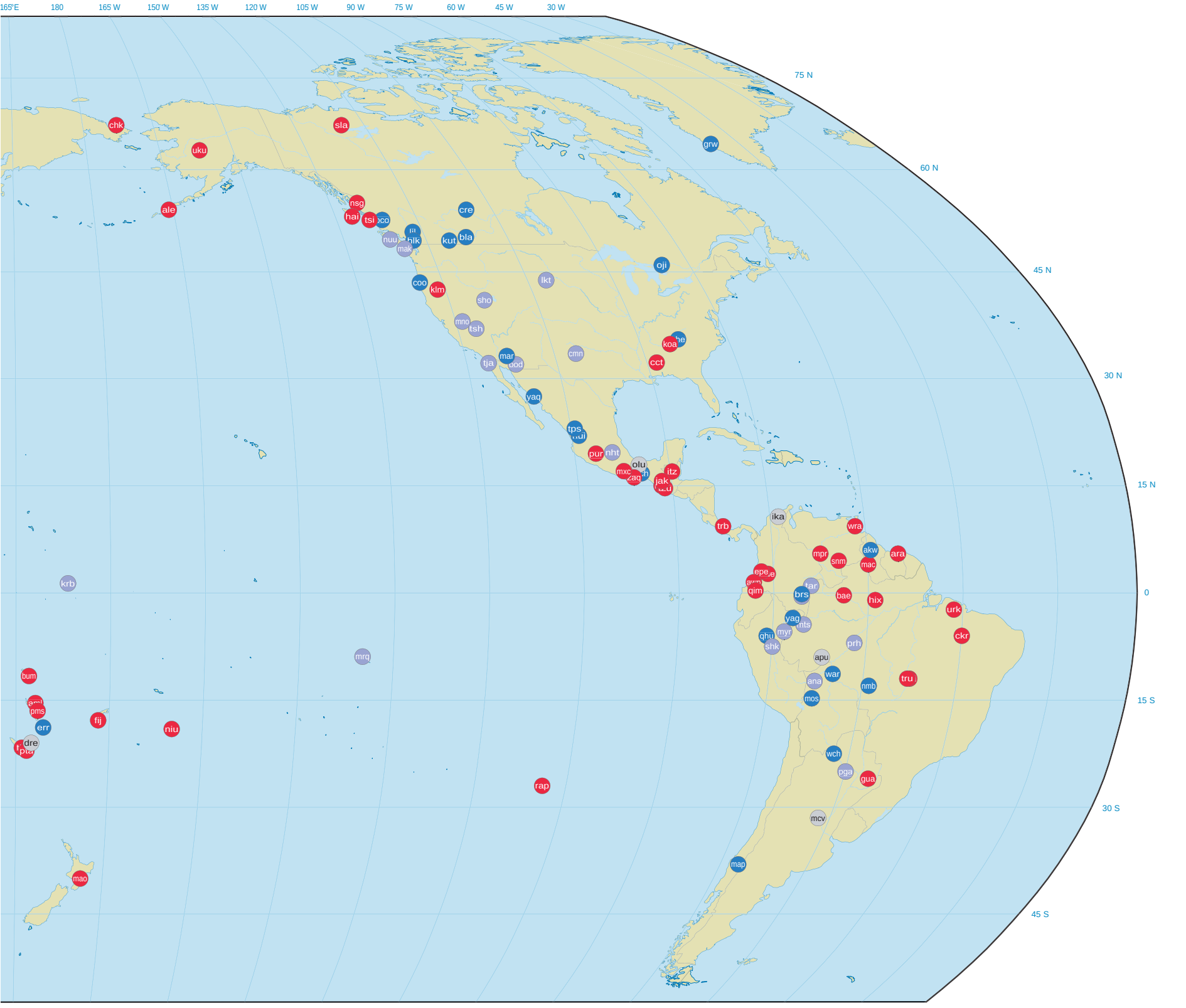


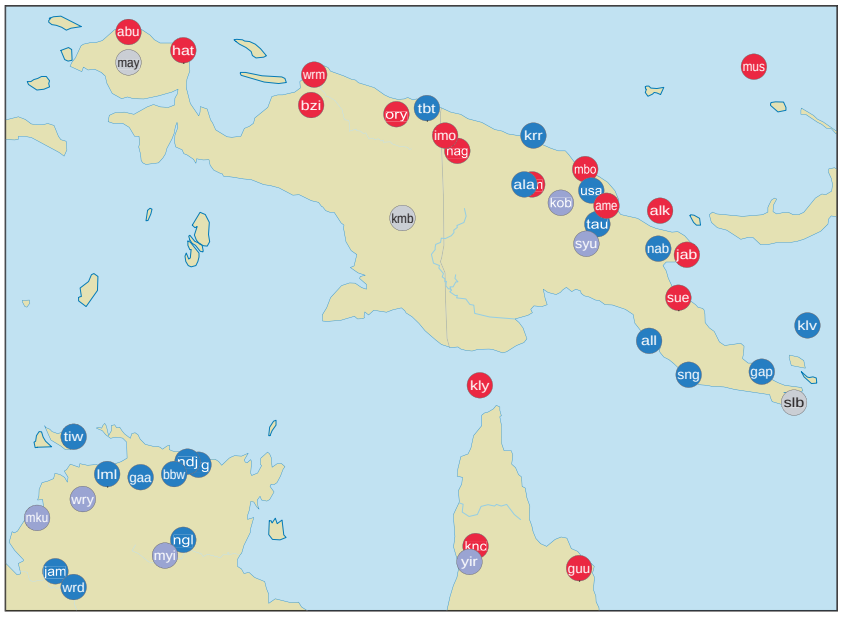


Nigeria and surrounding areas

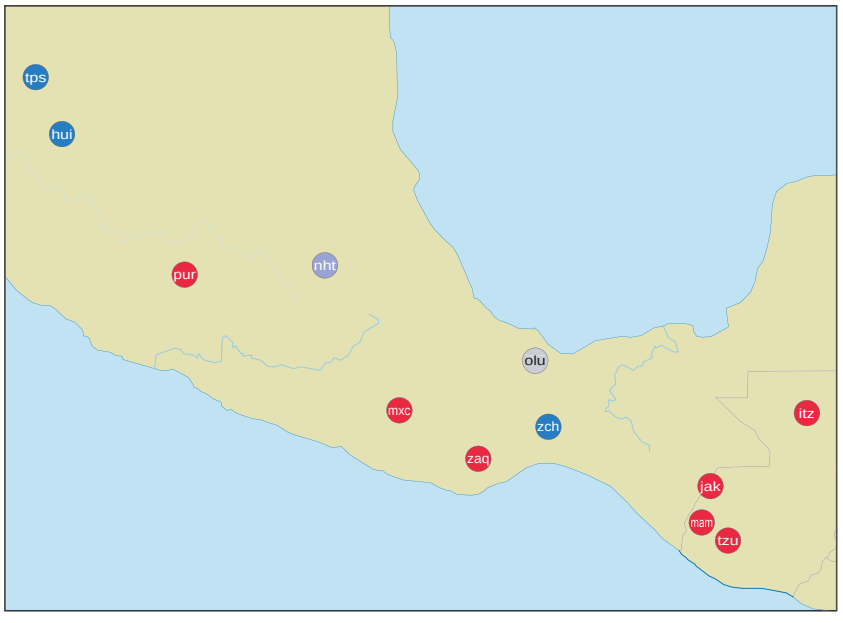


from Nepal to Thailand





New Guinea and northern Australia



Mesoamerica

106 Reciprocal Constructions

ELENA MASLOVA AND VLADIMIR P. NEDJALKOV

1 Introduction

Map 106 displays different ways in which languages encode reciprocal situations, i.e. situations like 'They love each other'. The reciprocal situation comprises at least two simple situations (e.g. *She loves him* and *He loves her*), so one way to encode such a situation is to combine expressions for two simple situations, as seen in (1) from Cantonese:

(1) Cantonese (Matthews and Yip 1994: 87)
Ngóh béi-mín kéuih kéuih béi-mín ngóh.
I give-face him he give-face me
'He and I respect each other.'

Apparently, this strategy of encoding reciprocated situations is possible in all languages, although the extent to which such a pattern is conventionalized in grammar and common in actual discourse differs from language to language. However, most languages also have simple or complex **reciprocal markers** (verbal affixes, pronouns, particles, adverbs, etc.), which, if combined with a verb within one clause, signal that the clause describes a reciprocal situation and not just the situation denoted by this verb, without repeating the verb for each simple situation. For example, English has two pronominal reciprocal markers, *each other* and *one another*; and in Kolyma Yukaghir (eastern Siberia), there is one reciprocal marker which is prefixed to the verb stem:

(2) Kolyma Yukaghir (own fieldwork)
a. *met tet-ul juo*
I you-ACC see(TR.1SG)
'I saw you.'
b. *mit* ***n'e-juo-ji:l'i***
we RECP-see-INTR.1PL
'We saw each other.'

English *each other* and Kolyma Yukaghir *n'e-* are **unambiguous**, i.e. they always express reciprocal meaning. In other languages, the reciprocal meaning may constitute one of several distinct functions of a grammatical construction. In Imbabura Quechua, for instance, the primary reciprocal marker is ambiguous between reciprocal and reflexive meanings, as seen in (3a); in addition, the reciprocal situation can be expressed by a suffix with *collective* meaning ('jointly, together'), as in (3b).

(3) Imbabura Quechua (Cole 1982: 91)
a. *wambra-kuna riku-**ri**-rka*
child-PL see-REFL/RECP-PST(3)
'The children saw each other.'
Or: 'The children saw themselves.'
b. *ñukanchi maka-**naju**-nchi*
we hit-COLL/RECP-1PL
'We hit jointly.' *Or*: 'We hit each other.'

This chapter focuses on one type of polysemy pattern that is associated with reciprocal constructions cross-linguistically, namely, reciprocal-reflexive polysemy as in (3a).

2 Definition of values

○	1. There are no non-iconic reciprocal constructions.	16
●	2. All reciprocal constructions are formally distinct from reflexive constructions.	99
◑	3. There are both reflexive and nonreflexive reciprocal constructions.	16
◔	4. The reciprocal and reflexive constructions are formally identical.	44
	total	175

2.1 The existence of non-iconic reciprocal constructions. In some languages, the encoding of a reciprocal situation always involves repetition of the main verb. This type of reciprocal encoding can be referred to as **iconic**, since the complex structure of the reciprocal situation is straightforwardly reflected in the structure of the grammatical construction. The iconic encoding can follow a conventionalized pattern distinct from the regular combination of two clauses. For example, the reciprocal construction in Amele (Madang, Trans-New Guinea; Papua New Guinea) involves third-person singular suffixes on each of the two forms of the verb, independently of the actual person of the reciprocal participants.

(4) Amele (Roberts 1987: 307)
*Ele ew-udo-**co-b*** *ew-udo-**co-b***
1DU despise-IO.3SG-DS-3SG despise-IO.3SG-DS-3SG
ow-a
1DU.SUBJ-PST
'We (two) despise each other.'

A particular case of iconic reciprocal encoding is reduplication of the verb stem, as in Godié (Kru, Niger-Congo; Côte d'Ivoire).

(5) Godié (Marchese 1986: 231)
wa wà-wà
they love-love
'They love each other.'

Languages that can express reciprocal meaning without repeating the verb, i.e. just by adding a reciprocal marker, show considerable variation in the range of contexts in which their reciprocal markers can be used. Most importantly, some reciprocal markers are compatible only with certain classes of verbs and can express a reciprocal relation only between certain participants in the situations signified by these verbs. For example, the reciprocal construction of West Greenlandic is available for transitive verbs only, and encodes a reciprocal relation between the core participants of the situation ("actor" and "patient"), as in the following example:

(6) West Greenlandic (Fortescue forthcoming)
***immi**-ssin-nut* *tuqun-niar-pusi*
REFL/RECP-2PL-ALL kill-FUT-2PL.IND
'You are going to kill each other/yourselves.'

Some reciprocal constructions are constrained to very small lexical classes of verbs; for example, Modern Hebrew has a reflexive/reciprocal construction which is used in its reciprocal meaning with about 10–15 verbs.

Secondly, there are languages where the reciprocal participants must be referred to by a single noun phrase (rather than by a conjunction of two or more noun phrases, as in *Mary and John*). This is the case in Ngiyambaa (Pama-Nyungan; Australia), where the reciprocal construction is marked by the verbal suffix *-la*.

(7) Ngiyambaa (Donaldson 1980: 166)
*miri gadha-**la**-nha*
dog(ABS) bite-RECP-PRES
'The dogs are biting each other.'

This construction cannot be used, however, if the reciprocal participants are denoted by two different noun phrases, as in *The dogs and the dingoes are biting each other*. Such a meaning can only be conveyed by a conjunction of two clauses, i.e. an iconic construction (Donaldson 1980: 163).

Thus, the availability of non-iconic reciprocal encoding in a given language can be viewed as a matter of degree, rather than a simple binary parameter. Since this study focuses on the polysemy of reciprocal markers (see §2.2), we have decided to take into consideration all non-iconic constructions, regardless of any language-specific

constraints. A language is assigned the feature value one (no non-iconic constructions) only if it has no such constructions at all. As a result, some languages can be assigned identical feature values (two, three, or four) in spite of considerable differences in the range of lexical or grammatical contexts in which specific non-iconic reciprocal constructions are available. The distinction between the values one, on the one hand, and two, three and four, on the other, is therefore of limited theoretical significance. Besides, since descriptive grammars differ in the amount of attention paid to lexically constrained or otherwise marginal constructions, the feature value one may have been mistakenly assigned to some languages that do have reciprocal markers. As a result, the number of languages without non-iconic reciprocal constructions may well have been overestimated.

*2.2 **Reciprocal-reflexive polysemy.*** Non-iconic reciprocal constructions fall into two types depending on whether or not they can also express reflexive meaning. This is a cross-linguistically significant parameter, since reciprocal and reflexive situations are clearly different (cf. 'They like each other' and 'They like themselves'), yet they share an essential feature which sets them apart from most other situation types: each participant plays two distinct roles in the same situation (e.g. respects somebody and is respected by somebody, as in (1)). The grammar of a language can reflect this similarity and use the same construction for reciprocal and reflexive situations, as illustrated by examples (3a), (6), and (8) from Wari' (Chapacura-Wanhan; Rondônia, Brazil).

(8) Wari' (Everett and Kern 1997: 191)
para *mana'* ***xujuhu'?***
therefore angry REFL/RECP.2PL
'Why are you angry with each other?' Or:
'Why are you angry with yourselves?'

Alternatively, reciprocal and reflexive constructions can be formally distinct (as in English, Kolyma Yukaghir (2b), or Ngiyambaa (7)). This parameter defines three language types: a language can have a reflexive reciprocal construction (as Wari', feature value four), a nonreflexive reciprocal construction (as in English, feature value two), or both (mixed type, as in German, feature value three).

If a language has a reflexive reciprocal construction, it does not mean that all or even most clauses with this marker are ambiguous. In Lithuanian, for instance, ambiguous reciprocal-reflexive clauses are impossible; although the language has a single marker that can express both meanings, only one meaning is possible for any given verb:

(9) Lithuanian (Geniušienė forthcoming)

a. *Petr-as* *ir* *On-a* *buèiuoja-**si**.*
Peter-NOM and Ann-NOM kiss-REFL/RECP
'Peter and Ann kiss each other.'

b. *Ona* *supa-**si**.*
Ann rock-REFL/RECP
'Ann is rocking [herself].'

Nevertheless, such constructions are classified here as reflexive reciprocals, along with those constructions that allow actual reflexive-reciprocal ambiguity. The reason for this decision is very simple: the semantics of a potentially ambiguous clause as a whole is bound to play an important role in its interpretation as either reflexive or reciprocal in any language, but the role of this factor in a non-native language is very difficult to assess, and this issue is rarely addressed in grammars.

3 Geographical distribution

Most languages in the sample for this study have non-iconic reciprocal constructions. The sample may somewhat magnify this tendency, since it was designed to include enough languages with non-iconic constructions to explore the distribution along the second parameter. However, this distortion is not likely to be very significant, since the *WALS* sample of 100 languages also contains a relatively small number of languages without non-iconic constructions (ca. 13 per cent); this type is not attested in South America, in Europe and in northern and Central Asia, and in Australia.

Turning now to languages that have non-iconic reciprocals, the map shows a clear distinction between two large areas, Eurasia and the rest of the world. Non-reflexive reciprocals are overwhelmingly dominant in Eurasia: the sample contains no Eurasian languages that have only reflexive reciprocals (value four) (this does not mean, of course, that there are no such languages, simply that this type is likely to be less frequent here). This large "nonreflexive" area stretches to the Pacific (excluding Australia), although the dominance pattern is less straightforward in the Pacific region (ca. 75 per cent of languages with reciprocals have nonreflexive reciprocal markers). The rest of the world has no dominant type, i.e. reflexive and nonreflexive reciprocals are more or less evenly distributed across major areas. This situation appears to reflect the diachronic instability of reciprocal types.

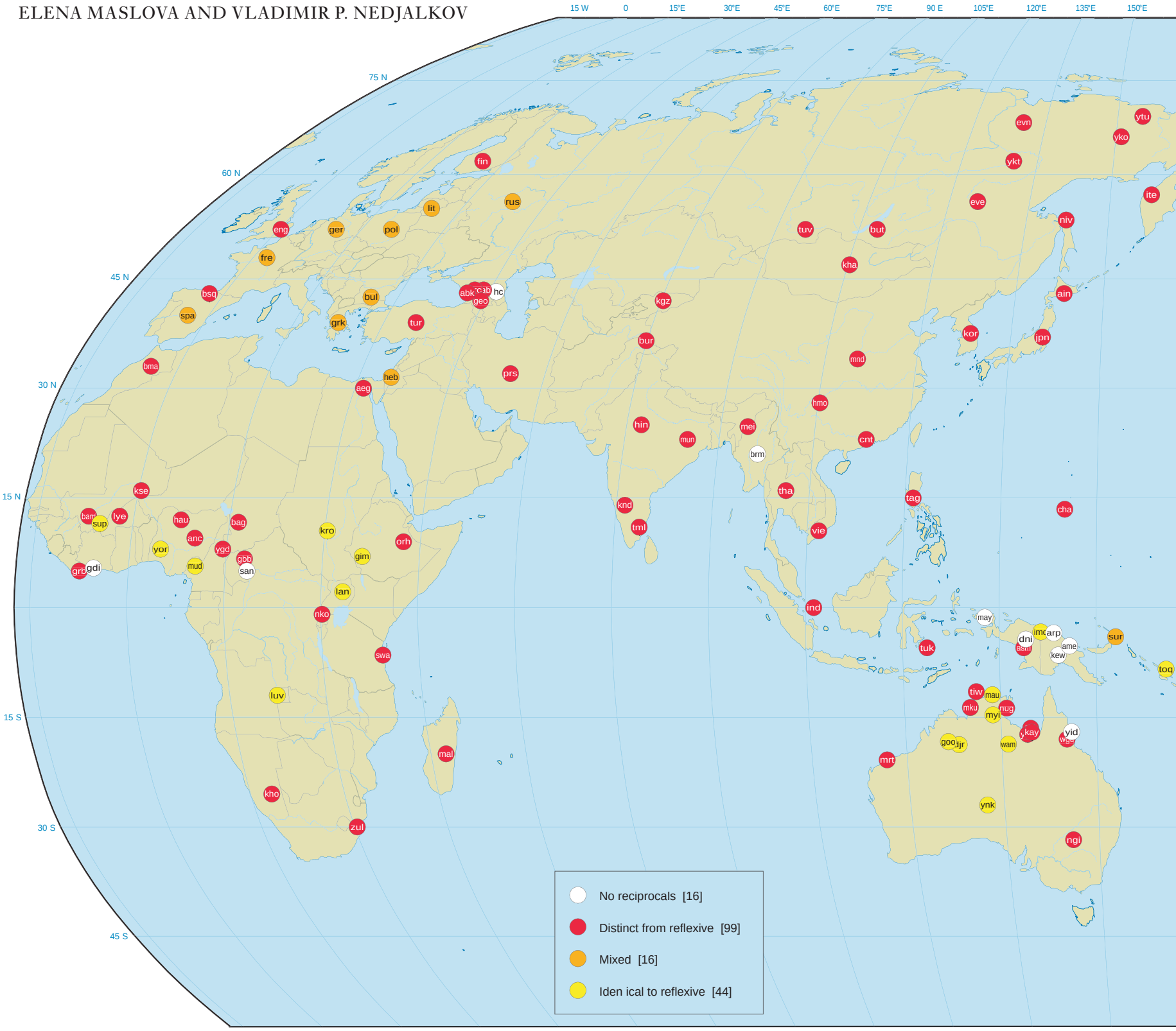
Another interesting pattern distinguishes Europe and the rest of the world: the mixed type (value three) is common in Europe (about half of the European languages in the sample) and very infrequent elsewhere: the 100-language sample contains only one such language outside Europe, Hixkaryana (Carib). This pattern may seem to be better accounted for in terms of common genetic origin, since almost all mixed-type European languages are Indo-European; on the other hand, Hindi (a non-European Indo-European language) has only a nonreflexive reciprocal, which means that common genetic origin cannot be the only determinant factor. It may be the case that the mixed reciprocal encoding should be considered as one of the several "exotic" traits of so-called "Standard Average European" (Dahl 1990, Haspelmath 2001).

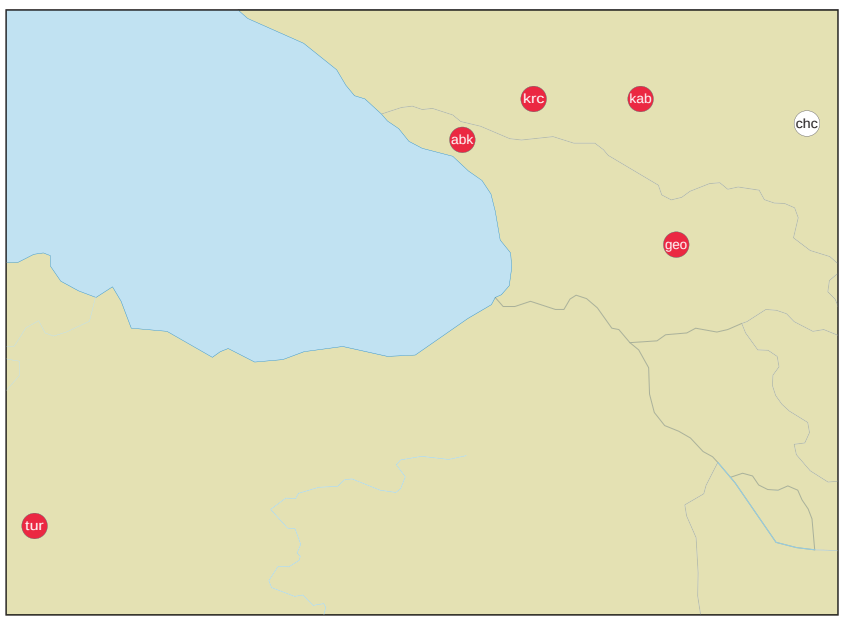
4 Theoretical issues

The reflexive-reciprocal polysemy pattern has been investigated within the frameworks of empirical typology (Geniušienė 1987), cognitive linguistics (Kemmer 1993), and grammaticalization theory (Heine 1999). The major theoretical issue within functionally oriented approaches has been the semantic affinity between reflexives and reciprocals which motivates cross-linguistically recurrent extension of prototypically reflexive constructions to cover reciprocal situations, and the resulting semantic bleaching of the reflexive markers. Nonreflexive polysemous reciprocals have been discussed in Lichtenberk (1985, 1999); Kemmer (1996); Maslova (1999). Detailed descriptions of various types of polysemy patterns involving reciprocals (including reflexive reciprocals) in a variety of languages can be found in (Nedjalkov et al. forthcoming). Within generative linguistics, reflexives and reciprocals used to be considered as instances of essentially the same syntactic phenomenon (anaphor); this approach has recently been challenged on the basis of both theoretical and empirical observations (see Everaert 1999 for an overview and discussion).

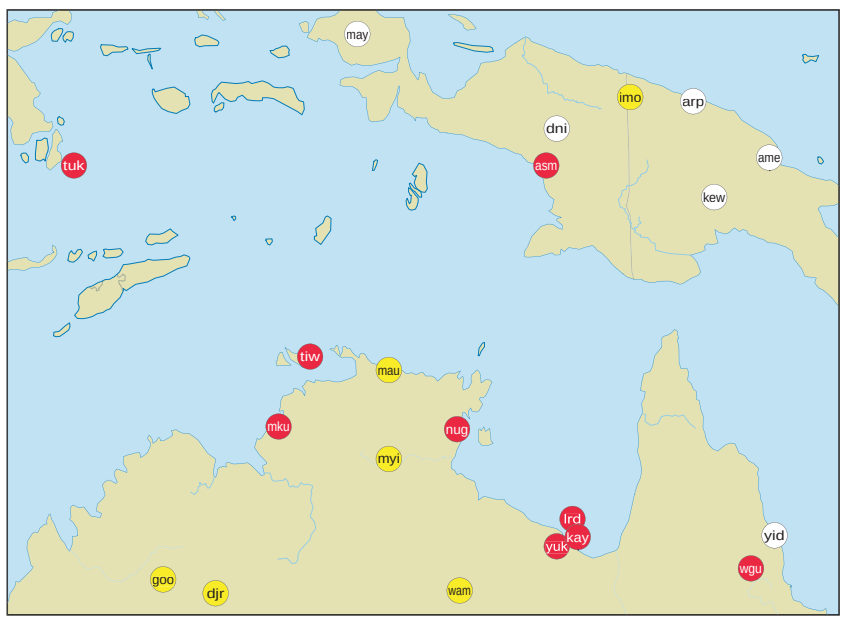
106 Reciprocal Constructions

ELENA MASLOVA AND VLADIMIR P. NEDJALKOV

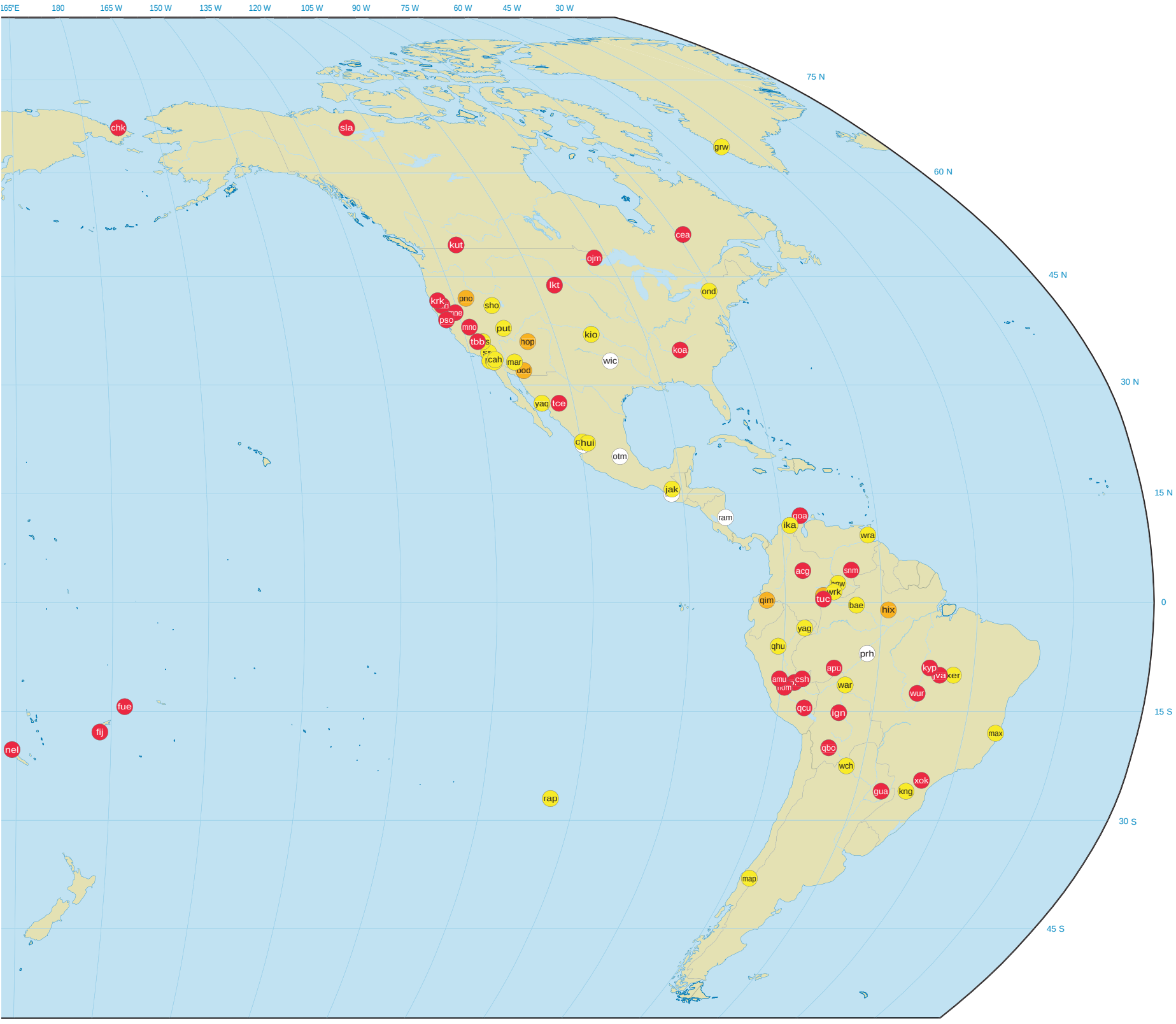


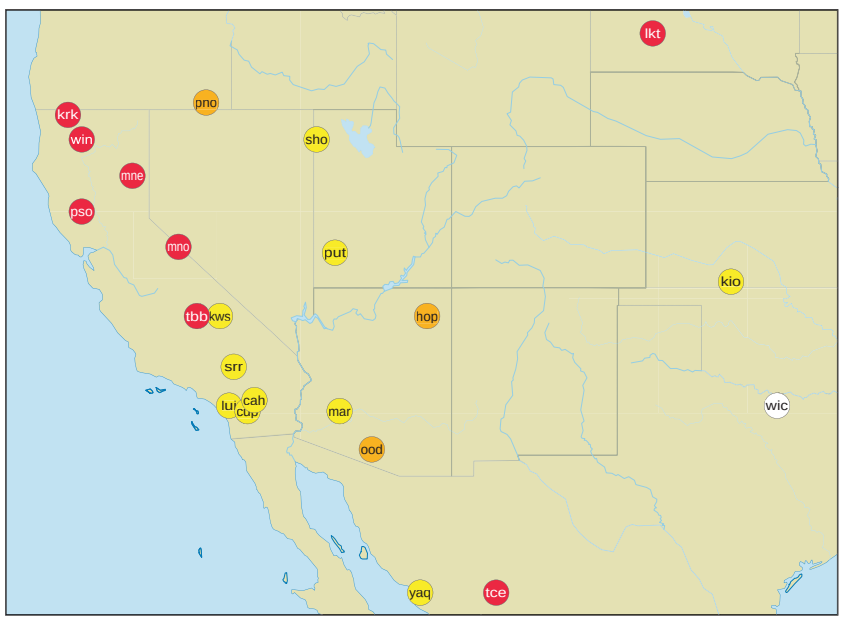


Caucasus and adjacent areas

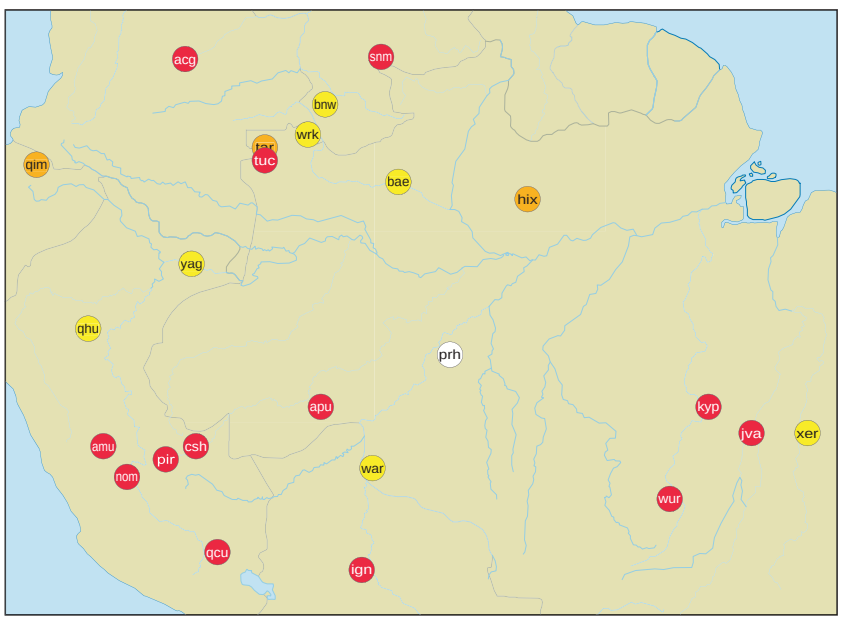


New Guinea and northern Australia





western North America



northwestern South America

107 Passive Constructions

ANNA SIEWIERSKA

1 Defining the values

Map 107 depicts the geographical distribution of passive constructions. Only two values are represented:

1.	There is a passive construction	162
2.	There is no passive construction	211
	total	373

A construction has been classified as **passive** if it displays the following five properties:

(i) it contrasts with another constuction, the **active**;

(ii) the subject of the active corresponds to a non-obligatory oblique phrase of the passive or is not overtly expressed;

(iii) the subject of the passive, if there is one, corresponds to the direct object of the active;

(iv) the construction is pragmatically restricted relative to the active;

(v) the construction displays some special morphological marking of the verb.

A prototypical example of the passive, as defined above, is given in (1b) from Swahili.

(1) Swahili (Ashton 1947: 224)

a. *Hamisi a-li-pik-a chakula*
Hamisi 3SG-PST-cook-IND food
'Hamisi cooked the/some food.'

b. *chakula ki-li-pik-w-a (na Hamisi)*
food 3SG-PST-cook-PASS-IND by Hamisi
'The food was cooked (by Hamisi).'

(1b) is an example of a **synthetic** passive, where the lexical verb exhibits some form of marking, here the suffix *-w*, which is absent from the active. Synthetic passives contrast with **periphrastic** or **analytical** passives, in which the special verbal morphology involves the use of a participial form of the lexical verb and an additional auxiliary verb, as in the English translation of (1b) and also in the Polish (2b).

(2) Polish (own knowledge)

a. *intensywne deszcze zniszczyły żniwa*
intensive rain destroyed harvest
'Intensive rain destroyed the harvest.'

b. *żniwa zostały zniszczone (przez intensywne deszcze)*
harvest remained destroyed by intensive rain
'The harvest was destroyed by intensive rain.'

In Swahili, Polish, and English the subject of the active may be expressed in the form of an oblique constituent or remain unexpressed. In many languages only the latter is possible: the subject of the active cannot be overtly present in the passive.

The three examples of passive clauses given so far are **personal passives**, i.e. passives with an overt lexical subject. Personal passives are typically seen as involving a process of agent demotion (from subject to oblique role or total suppression) and a process of patient promotion (from direct object to subject). There are also passive clauses which involve only agent demotion. These are called **impersonal passives**. An example of an impersonal passive is given in (3b) from Kannada (Dravidian; southern India).

(3) Kannada (Sridhar 1990: 215)

a. *ya:ro: i: nirNayav-annu khaNDisidaru*
someone this resolution-ACC denounce.PST.3PL.HUM
'Someone denounced this resolution.'

b. *i: nirNayav-annu khaNDisala:yitu*
this resolution-ACC denounce.INF.BECOME.3N
'This resolution was denounced.'

We see that the accusatively case-marked direct object *nirNayavannu* of the active (3a) retains its accusative case marking in the passive (3b). Moreover, the passive auxiliary *a:gu* 'become' is always in the third person singular neuter and thus shows no agreement with *nirNayavannu*. The direct object is thus not promoted to subject. This becomes even clearer when we compare the impersonal passive in (3b) with the Kannada personal passive in (4b), in which the direct object of the active appears in the nominative case and controls agreement with the passive auxiliary.

(4) Kannada (Sridhar 1990: 214)

a. *huDugaru ba:vuT-annu ha:risidaru*
boys flag-ACC fly.PST.3PL.HUM
'The boys flew the flag.'

b. *huDugar-inda ba:vuTa ha:risalpaTTitu*
boys-INSTR flag NOM fly.INF.PASS.PST.3SG.N
'The flag was flown by the boys.'

It is also important to note that in the Kannada impersonal passive, unlike the personal passive, it is not possible to express an overt agent. This, however, is not an integral feature of the impersonal passive. For instance, in Lithuanian, which also has both a personal and an impersonal passive, an overt agent can be expressed in both constructions. An example of the impersonal passive with an overt agent is given in (5b).

(5) Lithuanian (Ambrazas et al. 1997: 282)

a. *vaikaĩ miegójo sodè*
children.NOM slept.3PL garden.LOC
'The children slept in the garden.'

b. *vaikų̃ bùvo miẽgama sodè*
children.GEN be sleep.PRES.PART.N garden.LOC
'The children slept in the garden.' (lit. By the children was being slept in the garden.)

In both Kannada and Lithuanian the impersonal passive co-exists with the personal passive. This is also the case in Dutch, German, Hindi, Icelandic, Spanish, and Turkish. But there are languages which have only an impersonal passive, such as Kolami (Dravidian; Andhra Pradesh, India), Ute (Numic, Uto-Aztecan; Colorado), Tukang Besi (Western Malayo-Polynesian; Sulawesi, Indonesia), and Zuni (isolate; New Mexico). Languages with only impersonal passives have been classified here as exhibiting a passive on a par with languages with personal passives.

In languages which have no passive construction, agent demotion or suppression can be achieved by other means. Some languages simply allow the subject to be omitted. As shown in (6b), Awtuw (Sepik; Papua New Guinea) is such a language.

(6) Awtuw (Feldman 1986: 95)

a. *rey æye rokra-kay*
3SG M food cook-PERF
'He has cooked food.'

b. *æye rokra-kay*
food cook-PERF
'Someone has cooked food.'

In other languages, what would be expressed in English by an agentless passive is rendered by the use of an explicit impersonal or indefinite subject, such as the German *man* or French *on*, or simply the word for 'persons' or 'people' as in, for instance, Gude (Chadic, Afro-Asiatic; Nigeria and Cameroon).

(7) Gude (Hoskison 1983: 107)

kə	*digə*	*ənji*	*tə*	*ci*
COMP	beat.up	people	DEF	he

'He was beaten up.'

Still other languages achieve the same end by using the third-person singular or plural form of the verb. The latter is illustrated in (8) from Paamese (Oceanic; Vanuatu), in which the impersonal interpretation is confined to clauses lacking a corresponding third-person plural independent pronoun.

(8) Paamese (Crowley 1982: 180)

*(*kaile)*	*a-munumunu*	*Vauleli*
(*they)	3PL.REAL-drink	Vauleli

'There is drinking going on at Vauleli.'

There are several constructions which manifest some, but not all, of the five properties listed earlier as definitional for passive constructions, and which therefore have not been classified here as passives. First, there are **anticausative** (or *middle*, or *mediopassive*) constructions, as in (9b) from Gumawana (Oceanic; Papua New Guinea).

(9) Gumawana (Olson 1992: 349)

a. *boile* *iyana* *ka-kone-di*
yesterday fish 1PL.EXCL-trap.TR-3PL
'Yesterday we trapped many fish.'

b. *iyana* *bogina* *si-kona*
fish PERF 3PL-trap
'The fish are already trapped.'

Anticausative constructions resemble agentless passives in having a subject which is semantically a patient rather than an agent. However, in the anticausative, unlike the passive, there is no covert agent. The situation or event is depicted as being brought about spontaneously without the involvement of an agent. That this is indeed so is evinced by the fact that it is not possible to add to an anticausative construction an agentive manner adverb such as *deliberately* or *on purpose*. Compare the English passive (10a) with the anticausative (10b).

(10) a. *The door was opened deliberately.*
b. *The door opened (*deliberately).*

Second, there are constructions called **inverses** (see e.g. Cooreman 1987, Givón 1994). Inverse constructions are best known from the Algonquian languages, in which the direct voice is used if the agent is more topical or ontologically salient than the patient, and the inverse if the patient is more topical or ontologically salient than the agent. Traditionally the more topical or salient participant is called the *proximate* and the less salient or topical one the *obviative*. The direct/inverse opposition is illustrated in (11) from Plains Cree (Algonquian; Canada).

(11) Plains Cree (Wolfart 1973: 25)

a. *sēkih-ēw* *nāpēw* *atim-wa*
scare-DIR man.PROX dog-OBV
'The man scares the dog.'

b. *sēkih-ik* *nāpēw-a* *atim*
scare-INV man-OBV dog.PROX
'The man scares the dog.'

The inverse is similar to the passive in functional-pragmatic terms (Givón 1994). In both constructions the patient is more topical than the agent. However, whereas in the passive the agent is extremely nontopical or indeed simply suppressed, in the inverse the agent retains considerable topicality. Accordingly, the two constructions differ with respect to the properties of the agent. The agent in the passive, if expressed, is a syntactic adjunct. In the inverse, on the other hand, it is a syntactic argument. This is evinced by the obligatoriness of the agent in the inverse as opposed to the passive and by the ability of the agent of the inverse, for example, to determine verbal agreement or participate in various syntactic processes. The properties of the agent have therefore been used here as criterial for distinguishing the passive from the inverse.

Given the above, I have also not treated as passive the so-called non-actor focus constructions in the Philippine languages, illustrated in (12b) from Cebuano (Western Malayo-Polynesian; Philippines).

(12) Cebuano (Valkama 2000: 13–14)

a. *mo-palit* *ang* *tawo* *ug* *libro*
ACTOR.FOCUS-buy TOP man a book
'The man will buy a book.'

b. *palit-on* *sa* *tawo* *ang* *libro*
buy-GOAL.FOCUS the man TOP book
'The man will buy the book.'

There has been a long-standing controversy on whether or not the Philippine focus system should be considered to be a voice opposition and if so of what type—active/passive, ergative/antipassive, or even direct/inverse (see e.g. Siewierska 1984: 79–86 and the references cited there). The issue is a complex one and cannot be done justice to here. My main arguments against a passive analysis of non-actor focus clauses are: they exhibit a very high text frequency; the agent is typically overt and manifests some properties associated with syntactic arguments as opposed to adjuncts; they are semantically highly transitive in the sense of Hopper and Thompson (1980); and the verb does not exhibit *special* marking, since each of the various focus types, including actor focus, has its own dedicated verb marking. In sum, non-actor-focus clauses do not appear to be pragmatically restricted vis-à-vis their actor-focus counterparts.

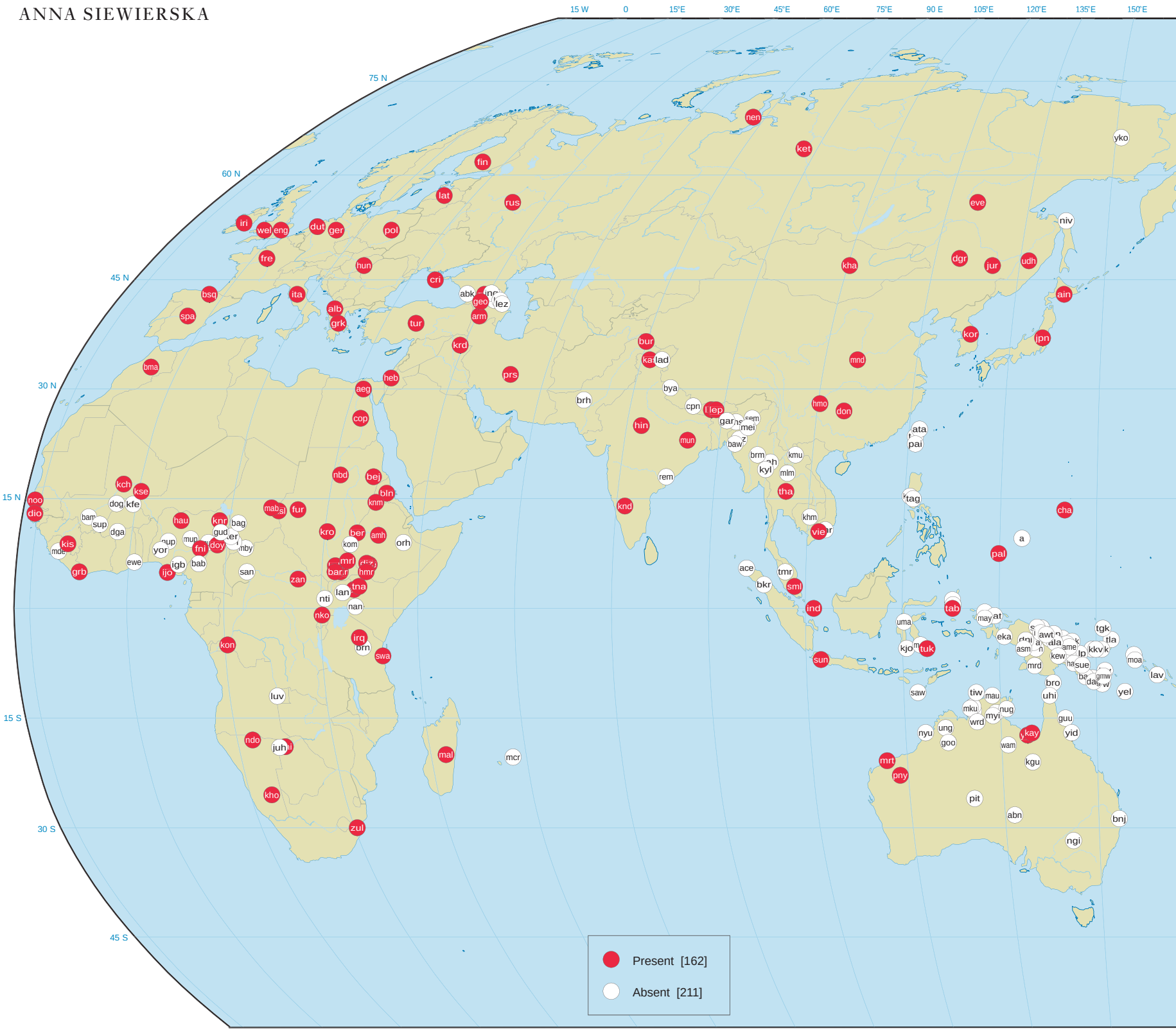
2 Geographical distribution

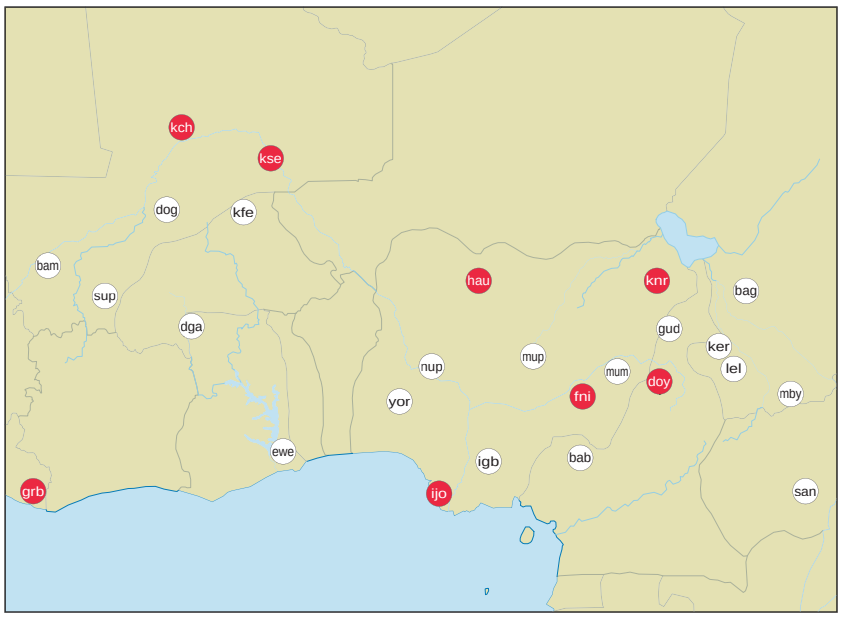
Passive constructions occur in 44 per cent of the languages in the sample. They are most common among the languages of Eurasia and Africa. They are also regularly found in the Americas, particularly North America. They are somewhat less frequent in South-East Asia and the Pacific. In Australia they are attested only in a couple of Tangkic languages spoken in the Gulf of Carpentaria and a few Ngayarda languages in southwestern Western Australia. In New Guinea they seem not to occur at all.

In Eurasia passives are frequent everywhere apart from the Caucasus and the Tibetan languages of India and Nepal. In Africa passives are highly common among the Nilo-Saharan languages, and only slightly less so in Afro-Asiatic. Of the Niger-Congo languages in the sample only about half display passive constructions. Passives are less frequent particularly around the coast of West Africa. In North America passives are found mainly in the western part of the continent. In South America they occur chiefly among the languages of the Amazon basin; they are particularly conspicuous by their absence among the languages along the west coast of the continent.

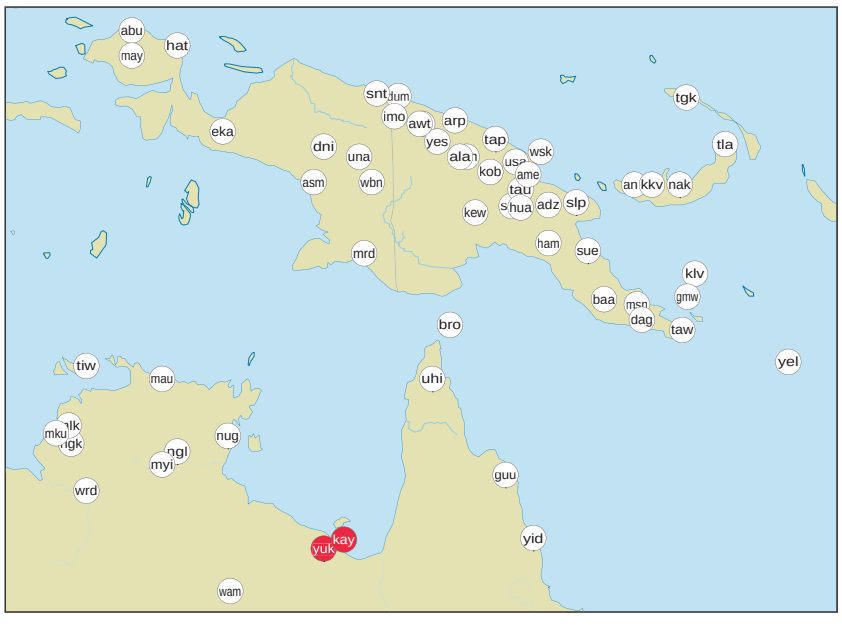
107 Passive Constructions

ANNA SIEWIERSKA

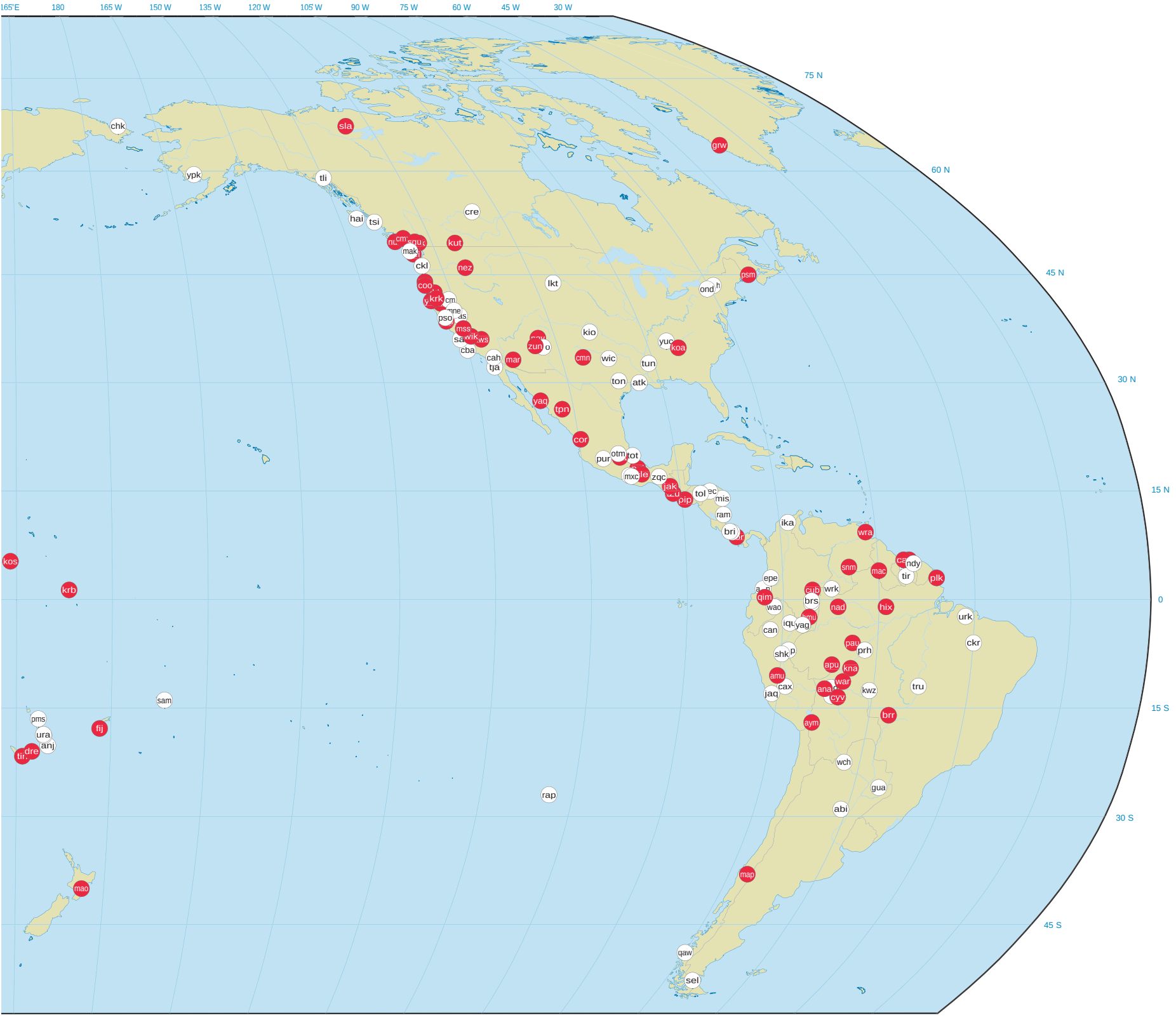




Nigeria and surrounding areas



New Guinea and northern Australia



western United States

southern Mexico and Central America

108 Antipassive Constructions

MARIA POLINSKY

1 Introduction

An antipassive construction is a derived detransitivized construction with a two-place predicate, related to a corresponding transitive construction whose predicate is the same lexical item. In the basic transitive construction, the patient-like argument is realized as a direct object; in the antipassive construction, that argument is either suppressed (left implicit) or realized as an oblique complement. The term **antipassive** (Silverstein 1972) was coined to indicate that the construction is a mirror image of the passive: in the passive, the suppressed or demoted argument is the agent-like argument, in the antipassive, the patient-like argument. An example of a transitive/antipassive alternation is given in (1a–b).

(1) Chukchi (Kozinsky et al. 1988: 652)

a. *ʔaaček-a* *kimitʔ-ən* *ne-nlʔetet-ən*
youth-ERG load-ABS 3PL.SUBJ-carry-AOR.3SG.OBJ
'The young men carried away the/a load.' (transitive)

b. *ʔaaček-ət* ***ine-nlʔetet-gʔe-t*** *kimitʔ-e*
youth-ABS ANTIP-carry-AOR.3SG.SUBJ-PL load-INSTR
'The young men carried away the/a load.' (antipassive)

In (1a), the transitive verb 'carry' agrees with the ergative subject and absolutive object. In (1b), the verb is marked with the antipassive prefix *ine-* and no longer agrees with the object; the object is now expressed by an oblique case (instrumental).

A verb in the antipassive is derived from the corresponding transitive verb, often with the help of overt morphology. For example, in (1b), the verb *nlʔetet-* bears the dedicated antipassive prefix *ine-*. Languages also make use of antipassive markers that are syncretic with other categories, thus functioning simultaneously as (for example) detransitivizers or aspect/modality markers.

Some languages show syncretism between the morphology of the antipassive and the morphology of other detransitivizing operations, most commonly reflexivization (e.g. in Pama-Nyungan languages of Australia). The antipassive verb can take on other formal characteristics of intransitive verbs; one such common characteristic is the change in the agreement pattern, as illustrated by the Chukchi examples above. In some languages (for example, in Mayan: England 1983b; 1988), a change in verbal agreement may be the sole indication of the antipassive. Finally, there is a correlation between the antipassive and the imperfective/irrealis (see §4.3).

2 Definition of values

Given that our knowledge of variation in antipassive constructions is far from complete, it is possible that some languages which are identified on the map as having no antipassive will ultimately be reanalysed as having an antipassive construction.

Among the languages which have the antipassive construction, the major division is between the type that requires the patient-like argument to be unexpressed (the **implicit argument type**) and the type where the patient-like argument can be expressed by an **oblique complement**. The following values are represented on the map:

1.	Antipassive with patient-like argument left implicit	18
2.	Antipassive with patient-like argument expressed as oblique complement	30
3.	No antipassive	146
	total	194

As shown on the main map, the antipassive is found across different language families and geographical areas; the antipassive with an oblique complement seems to be more widespread than the antipassive with an implicit argument.

3 The antipassive construction and ergativity

One of the vexing questions concerning the distribution of the antipassive is its correlation with **ergativity** (see Chapters 98–100 for ergative constructions). Some authors insist on the link between the antipassive and ergativity (Silverstein 1976; Dixon 1979; Spencer 1991: 24), while others propose that the antipassive is not limited to ergative languages (Heath 1976; Postal 1977; Davies 1984; Givón 1984; Lidz 1996). Table 1 lists all the languages of the sample with antipassives; it shows no principled correlation between ergativity and the antipassive. The transitive/antipassive alternation is more visible in an ergative language, where it typically involves a change in subject case marking from ergative to absolutive, as in (1a–b) above.

Table 1 The antipassive construction and case marking

Accusative	Ergative
Acoma, Cahuilla, Canela-Krahô, Chamorro, Choctaw, Comanche, Cree, Kiowa, Koyraboro Senni, Krongo, Lango, Lavukaleve, Nez Perce, Ojibwa, Paiwan, Sanuma, Thompson	Archi, Bezhta, Cakchiquel, Central Yup'ik, Chechen, Chukchi, Copainalá Zoque, Diyari, Djaru, Dyirbal, Embaloh, Godoberi, Gooniyandi, Halkomelem, Hunzib, Jakaltek, Kabardian, Kapampangan, Lai, Lak, Mam, Mangarrayi, Päri, Tsez, Tzutujil, Wardaman, Warrungu, West Greenlandic, Yidiny, Yukulta

4 Theoretical issues

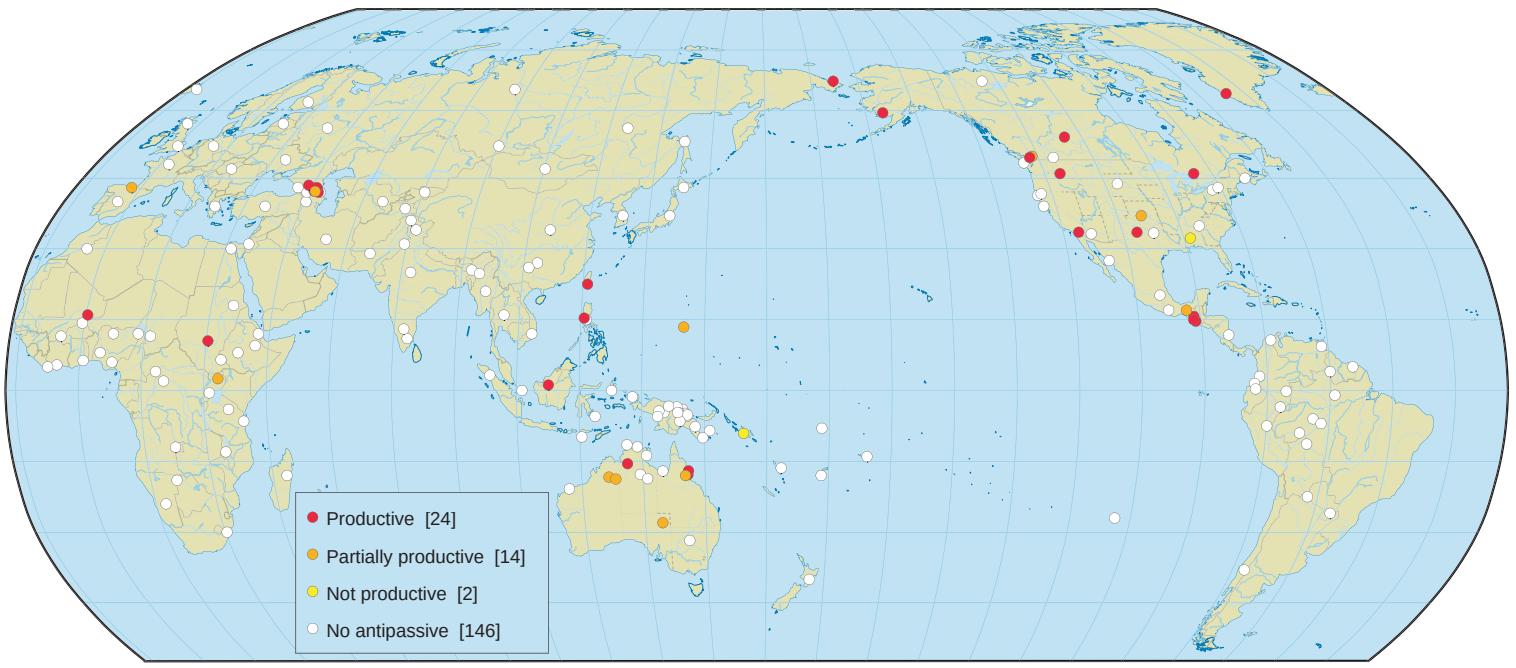
4.1 Productivity of the antipassive construction. Antipassives often apply only to a certain subset of transitive predicates (Gibson 1980; Chung 1998: 39). Productive antipassives show genealogical and areal clustering, as indicated on the inset map: they are found in some Mayan, Salishan, Nakh-Daghestanian, Austronesian, and Australian languages.

1.	Productive	24
2.	Partially productive	14
3.	Not productive	2
4.	No antipassive	146
	total	186

Values of Map 108A Productivity of the Antipassive Construction

4.2 Morphological marking. The variation observed with respect to the marking of antipassives has to do with whether or not the language has a **dedicated antipassive marker**. A dedicated marker can often be associated diachronically with other functions (e.g. middle, reflexive), so what matters is its nonsyncretic synchronic status. In our sample, languages with a dedicated antipassive marker include: Acoma, Chamorro, Chukchi, Comanche, Diyari, Dyirbal, Godoberi, Gooniyandi, West Greenlandic, Halkomelem, Hunzib, Jakaltek, Kiowa ("Type 2" antipassive, see Watkins 1984: 140–1, 149–50), Koyraboro Senni, Krongo, Lango, Mam, Eastern Ojibwa, Tzutujil, Warrungu, Copainalá Zoque.

In languages that have several different antipassive markers, these markers can co-occur (or "stack"). Stacking is found in Chukchi (the productive antipassive prefix *ine-* and the semiproductive antipassive suffix *-tko-*, Kozinsky et al. 1988: 661) and in Halkomelem (the middle antipassive suffix *-əm/ -eʔəm-* and the activity suffix



Map 108A Productivity of the Antipassive Construction

-els-, Gerdts and Hukari forthcoming). The semantic contrast between the stacked and the single antipassive is, however, unclear.

4.3 Semantic and discourse functions of the antipassive. The semantic import of the antipassive has to do with the **affectedness** and **individuation** of the patient.

The use of a prototypical transitive verb entails that the event denoted by that verb causes a change of state in the object participant (Tsunoda 1981; Hopper and Thompson 1980; Van Valin 1991; Dowty 1991, among many others). The semantic function of the antipassive is to cancel such an entailment; this correlates with the marking of the patient as an oblique complement. Comparable effects can be found in the English **conative alternation**, whereby the referent of a direct object is construed as affected while the referent of a prepositional complement is not (Levin 1993: 5–11):

(2) a. *The hunter shot the bear* (entailment: the bear died)
b. *The hunter shot at the bear* (the bear may not be affected at all)

If there is no affected participant which allows one to measure out the effects of the event (the **incremental theme**, Dowty 1991), the event itself is interpreted as incomplete. This accounts for the high correlation between the use of the antipassive and the habitual, durative, iterative, and imperfective (Tchekhoff 1987; Cooreman 1994; Dixon 1994; Dowty 1991; van den Berg 2001: 60). A correlation between the use of antipassive and irrealis is found in Yukulta (Keen 1983).

Compared to the transitive construction, the patient participant in the antipassive construction is lower in individuation (or identifiability). A lower degree of individuation may be manifested by a number of structural possibilities, summarized in Table 2.

Table 2 Low individuation of the object in the antipassive

Antipassive preferred or required if the object is:	**Attested in**
Plural	Bezhta (Kibrik 1981)
Indefinite	West Greenlandic (Kalmár 1979; Seiler 1978; Bittner 1987)
Nonspecific	Archi (Kibrik et al. 1977)
Generic	Diyari (Austin 1981)
Implicit argument	Mayan languages (Smith-Stark 1978; England 1988)

A referent with low individuation is less likely to be maintained in discourse; this explains why the antipassive is often used to introduce episodic referents (Polinskaja and Nedjalkov 1987; Bittner 1987), whose identification is not crucial for communication.

However, there are languages (e.g. Yukulta) where the antipassive is required when the object is expressed by a first- or second-person pronoun, which are arguably the highest on the scale of individuation. Such a subtype contrasts with the subtype exemplified by Halkomelem, where the antipassive is impossible with a first- or second-person object (Gerdts 1988: 157). This variation supports the conclusion, made by several researchers, that the semantic and discourse functions of antipassives can and do differ across languages (Comrie 1978; Heath 1976; Cooreman 1988, 1994).

4.4 Structural functions of the antipassive. In a number of languages, only a particular grammatical function (subject, object, etc.) or case form can serve as the syntactic pivot for extraction, reference tracking across clauses, and other grammatical processes. Antipassivization makes the sole argument of the detransitivized verb accessible to relevant grammatical processes; in many ergative languages, where the absolutive is the pivot for all or most grammatical processes, antipassivization allows the subject to occur in the absolutive case. To illustrate the syntactic function of antipassives, let us turn to more Chukchi examples. In Chukchi, relativization is possible only for absolutive arguments (Polinsky 1994); thus, the subject of intransitives and the direct object can relativize directly, as shown by (3) for the subject absolutive. In order to relativize the subject of a transitive (4), it first needs to be converted into an absolutive, via antipassivization (5a); the subject of the antipassive is then relativized, as in (5b).

(3) a. *ŋinqey* *pəkir-gʔi* ‘The boy arrived.’
boy.ABS arrive-AOR.3SG
b. *[pəkərə-lʔ-ən]* *ŋinqey* ‘the boy that arrived’
arrive-PTCP-ABS boy

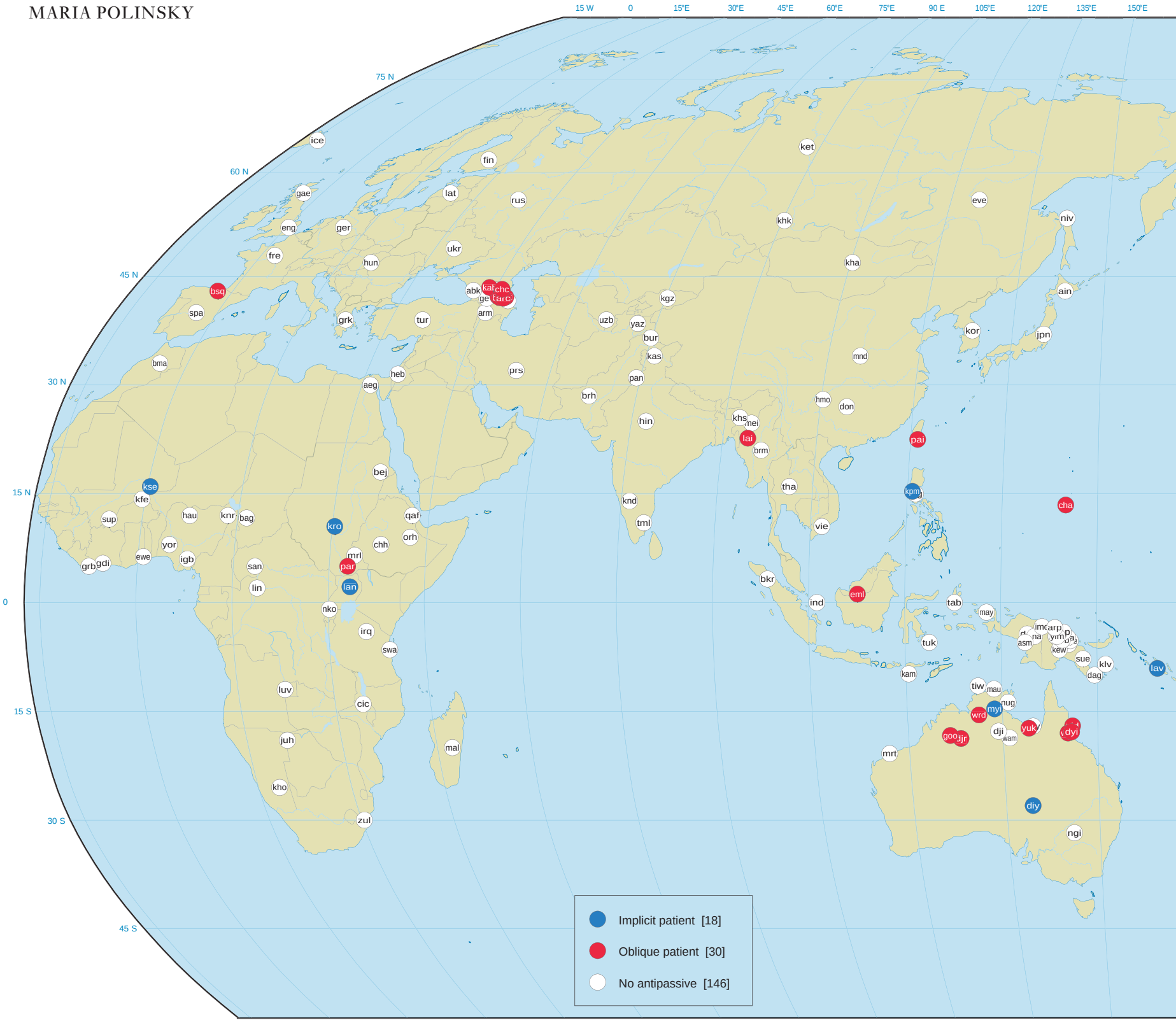
(4) a. *tumg-e* *ŋinqey* *rəyegtetew-nin*
friend-ERG boy.ABS save-AOR.3SG.3SG
‘The friend saved the boy.’
b. * [*ŋinqey* *rəyagtala-lʔ-ən*] *tumgətum*
boy.ABS save-PTCP-ABS friend
(‘the friend that saved the boy’)

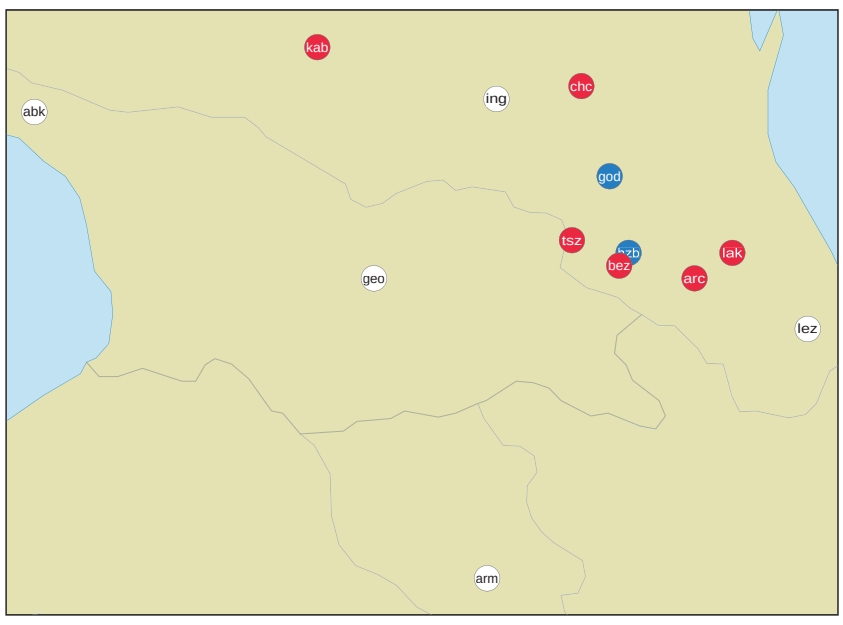
(5) a. *tumgətum* *ŋinqey-ək* *ine-nyegtele-gʔi*
friend.ABS boy-LOC ANTI-save-AOR.3SG
‘The friend saved the boy.’
b. [*ŋinqey-ək* *ine-nyegtelewə-lʔ-ən*] *tumgətum*
boy-LOC ANTI-save-PTCP-ABS friend
‘the friend that saved the boy’

The antipassive also establishes unambiguous subject-subject coreference across clauses; such use of the antipassive is observed in Australian languages (Tsunoda 1988; Cooreman 1994), Chukchi (Kozinsky et al. 1988: 689–97), and Mayan (England 1983b).

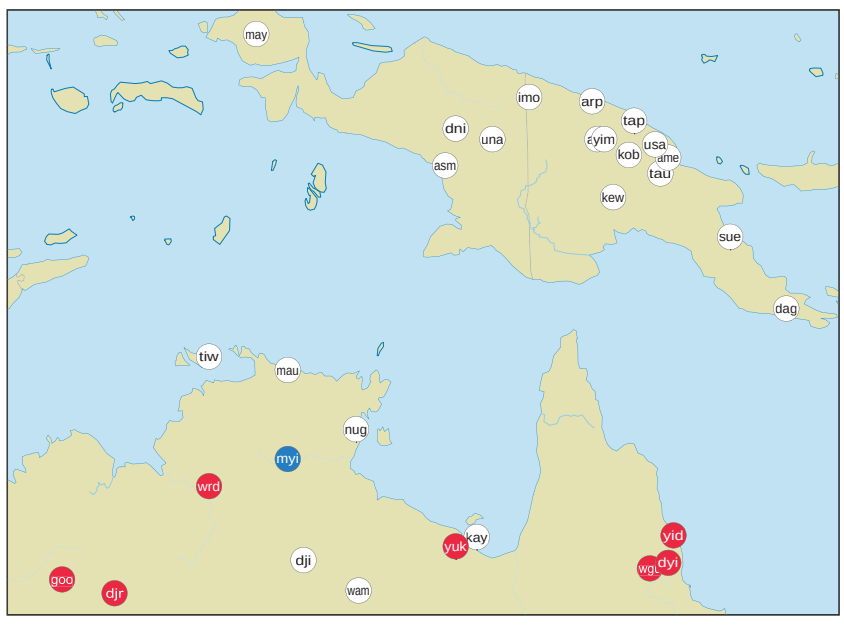
108 Antipassive Constructions

MARIA POLINSKY

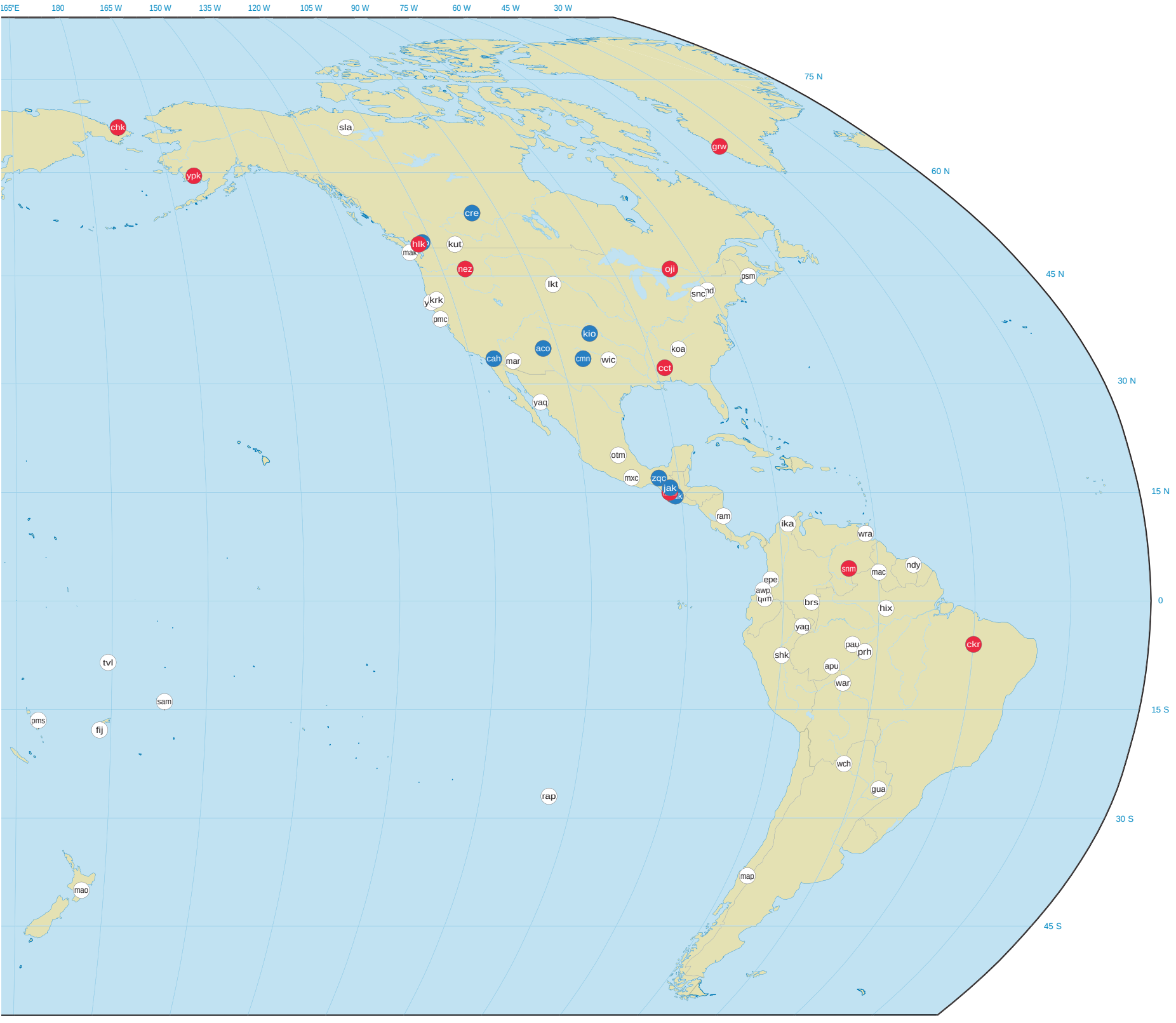


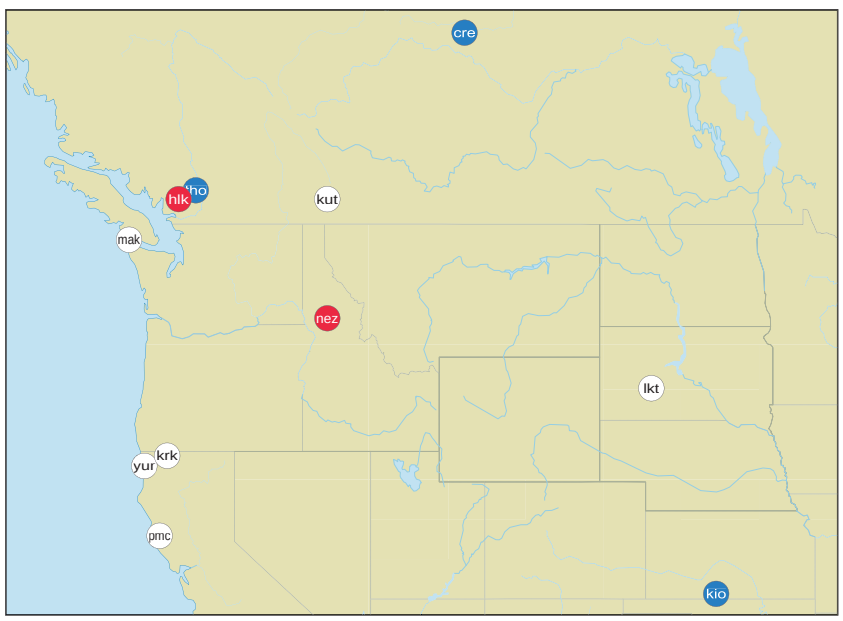


Caucasus

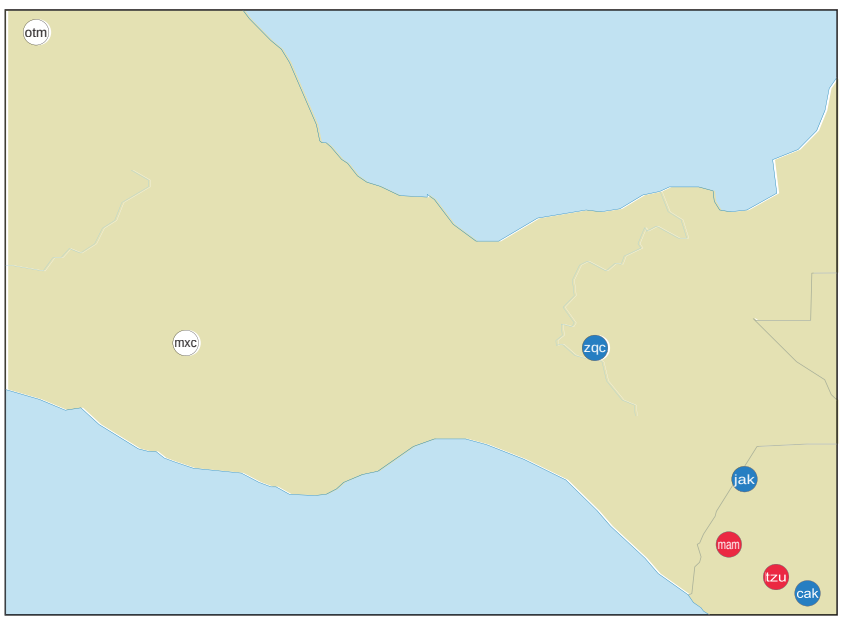


New Guinea and northern Australia





western North America



Mesoamerica

109 Applicative Constructions

MARIA POLINSKY

1 Introduction

In an **applicative construction**, the number of object arguments selected by the predicate is increased by one with respect to the basic construction. For example, in Tukang Besi (Austronesian; Sulawesi, Indonesia), the verb 'fetch' takes one theme object in the basic construction (as shown in 1a), but with the applicative marker it takes two objects, theme and benefactive (as shown in 1b).

(1) Tukang Besi (Donohue 1999: 231)

a. Basic construction, two-place predicate
no-ala *te* *kau*
3.REALIS-fetch the wood
'She fetched the wood.'

b. Applicative construction, three-place predicate
no-ala-ako *te* *ina-su* *te* *kau*
3.REALIS-fetch-APPL the mother-my the wood
'She fetched the wood (as a favour) for my mother.'

The term "applicative" is used to denote either the applicative construction or the verb in that construction. The term was first used by early Spanish missionary grammarians in the description of Native American languages, in particular Uto-Aztecan; it was later adopted by Bantuists and is now used for similar constructions all over the world. It is customary to restrict the designation "applicative" to those cases where the addition of an object is overtly marked on the predicate. Thus English pairs such as *She baked a cake—She baked Oscar a cake* do not count as basic-applicative alternation. The object added in the applicative construction is referred to as the **applied object**; if a verb without an applicative suffix takes an object, this latter object is referred to as a **basic object**. The verb from which the applicative is formed is called the **base**.

An applicative construction is a particular instance of a **double object construction**. Both constructions share the property that the two objects do not have to be distinguished by agreement morphology or case marking (see also Chapter 105). However, the primary use of the double object construction is for the encoding of the event of transfer. The presence of an applicative construction per se often entails that a nonderived double object construction is also present. But the opposite is not the case; for example, Dutch and English have the double object construction but not the applicative construction. Some dialects of Tamil and some Australian languages seem to be an exception in having the applicative/causative but no double object construction (Polinsky 1995; Austin 1997; Tsunoda 1998).

2 Definition of values

Two main parameters in which applicatives vary cross-linguistically include the transitivity of the base and the semantic role of the applied object. The main map for this chapter reflects the values of these parameters.

With respect to the **transitivity of the base**, the main distinctions are between applicatives formed (i) from a **transitive base only**, (ii) from an **intransitive base only**, and (iii) from **both bases** (these distinctions are shown by different shapes of the symbols on the map). There are also constraints on the number of arguments of the base verb; they are not shown in the map because they are harder to determine on the basis of grammars.

With respect to the **semantic roles of the applied object**, the most common role of the applied object is that of benefactive. Accordingly, the map differentiates applicatives whose applied object (i) is **limited to benefactive**; (ii) corresponds to the **benefactive and some other roles**; (iii) corresponds to **other roles to the exclusion of the benefactive**. These distinctions are shown by the different colours of the symbols on the map.

Of nine logically possible values (three different shapes, three different colours), seven are actually attested. In addition, there are of course many languages without applicatives. The map thus shows the following eight values:

1. Benefactive object only; both bases	16
2. Benefactive object only; transitive base only	4
3. Benefactive and other; both bases	49
4. Benefactive and other; transitive base only	2
5. Nonbenefactive object only; both bases	9
6. Nonbenefactive object only; transitive base only	1
7. Nonbenefactive object only; intransitive base only	2
8. No applicative construction	100
total	183

3 Geographical distribution

Applicatives are common in three geographical areas: Africa (mostly in Bantu), the western Pacific region (Austronesian), and North and Meso-America (Salish, Mayan, Uto-Aztecan). The main generalization seems to be that applicatives are commonly found in those languages that have little or no case marking of noun phrases in a clause and that have sufficiently rich verbal morphology to mark applicative formation on the predicate. The dearth of applicatives in Eurasia may thus be due to the widespread presence of rich nominal morphology in the languages of that area, and indeed, where applicatives are found is in languages like Abkhaz and Abaza (North-West Caucasian; Georgia and Russia), which have little or no overt case marking of noun phrases.

4 Other semantic roles (inset map)

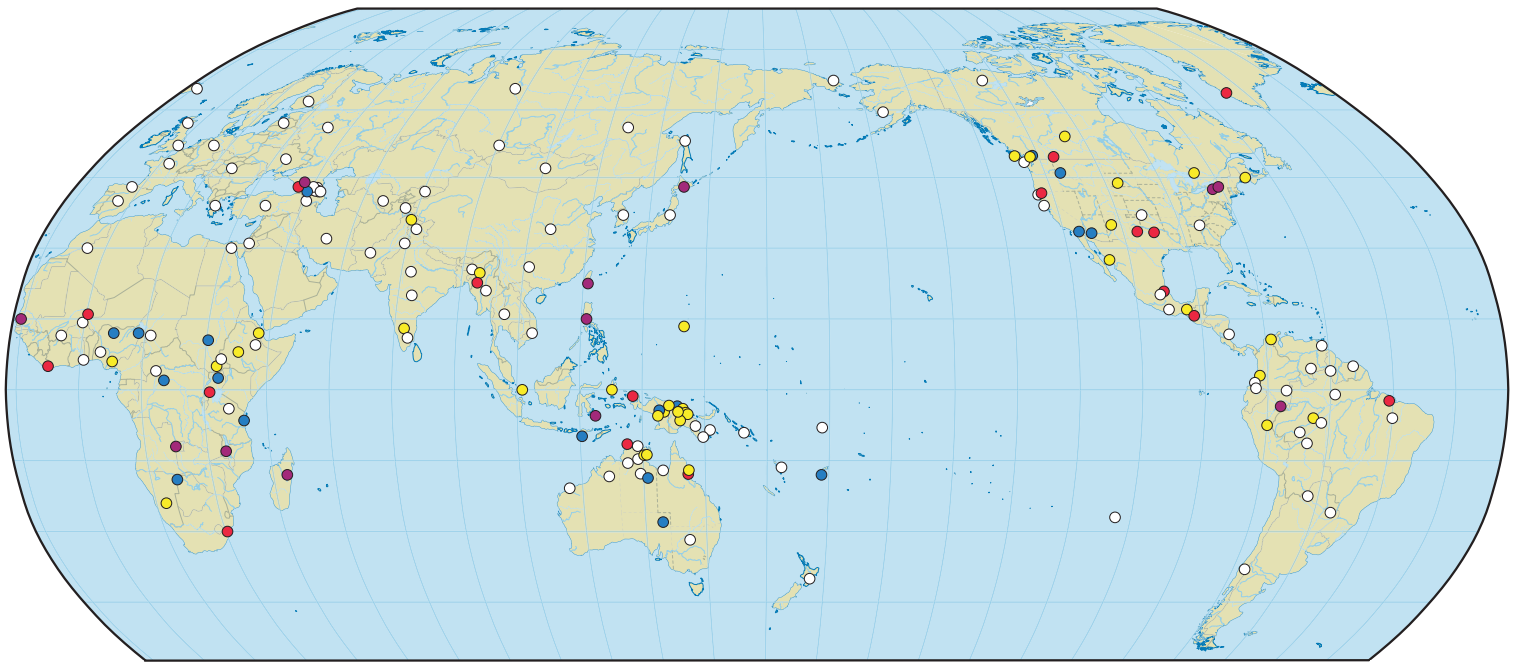
The main map in the atlas reflects the distribution of applicatives with respect to the semantic role benefactive, clearly the most common semantic role of the applied object. Other common semantic roles include **location** and **instrument**. The geographical distribution of these roles is shown on the inset map.

1. Instrument	17
2. Locative	18
3. Instrument or locative	12
4. No other roles (= only benefactive)	36
5. No applicative construction	100
total	183

Values of Map 109A Other Roles of Applied Objects

Additional semantic functions that may be associated with the applied object include possessor, circumstance/event (time), comitative, and substitute (a participant on whose behalf the action is performed). Comitatives and substitutives are quite common in applicatives of intransitives, e.g. in Lai (2) and in Kinyarwanda (3):

(2) Lai (Peterson 1999: 58)
ʔa-ka-Than-pii
3SG.SUBJ-1SG.OBJ-grow.up-APPL.COM
'He grew up with me.' [comitative]



Map 109A Other Roles of Applied Objects

(3) Kinyarwanda (own field notes)
umugabo *a-ra-geend-er-a* *umugóre*
man 3SG-PRES-travel-APPL-ASP woman
'The man is travelling instead/on behalf of the woman.'
[substitute]

It is somewhat puzzling that, in quite a few languages of Australia, comitative applicatives can be formed from intransitives, but neither comitative nor benefactive applicatives can be formed from transitives. Such comitative applicatives are sometimes marked by the same morphology as instrumental applicatives formed from transitives. It is possible that these Australian-type applicatives and applicatives elsewhere represent different phenomena.

5 Theoretical issues

Applicative verbs and constructions have generated a significant amount of research dealing with the morphosyntax of applicatives, transitivity, VP structure, distinct object relations, argument structure, and mapping from argument to syntactic structure. For an overview, see Alsina 1996, Alsina and Mchombo 1993, Peterson 1999.

5.1 Transitivity of the base. The intransitive base of applicatives is less common than the transitive base. This is quite clear from the map, and there are only two languages in the sample that form applicatives from the intransitive base exclusively (Fijian, Wambaya). The overall tendency is that if a language has applicatives formed from the intransitive base, it also has applicatives formed from the transitive base.

The following explanation may be offered as to why the applicative is uncommon with intransitives. Adding an object to an intransitive amounts to creating a transitive verb. In order to transitivize an intransitive, languages typically use causativization, thus increasing the complexity of the event structure (CAUSE(V)); with the applicative, the event structure is not modified, just another participant is added. Under causativization, the argument added to the argument structure of the verb is agent; under non-causative transitivization, the added argument is theme (or patient). Applicative formation results in adding arguments other than agent and theme, and may thus be constrained by the general hierarchy of semantic roles: agent > theme (patient) > goal (recipient, benefactive) > location > other.

Although the constraint on applicative formation from intransitives seems not to be absolute, a particular subset of intransitives, namely unaccusative predicates (those whose subject originates as an object in the underlying structure), resist applicativization (Baker 1998). However, even this generalization does not hold in some languages, for instance, in Halkomelem (Gerdts 1988), Lai (Peterson 1999), and Sesotho (Machobane 1989), suggesting that it is just a strong tendency.

5.2 Morphosyntax. With respect to the formation of the applicative predicate, much discussion has centred on the question of which object, basic or applied, is structurally superior. Some researchers have proposed that the two objects may be structurally equal (Gary and Keenan 1977; Seiter 1980), but most analyses place the base object higher or lower than the applied object.

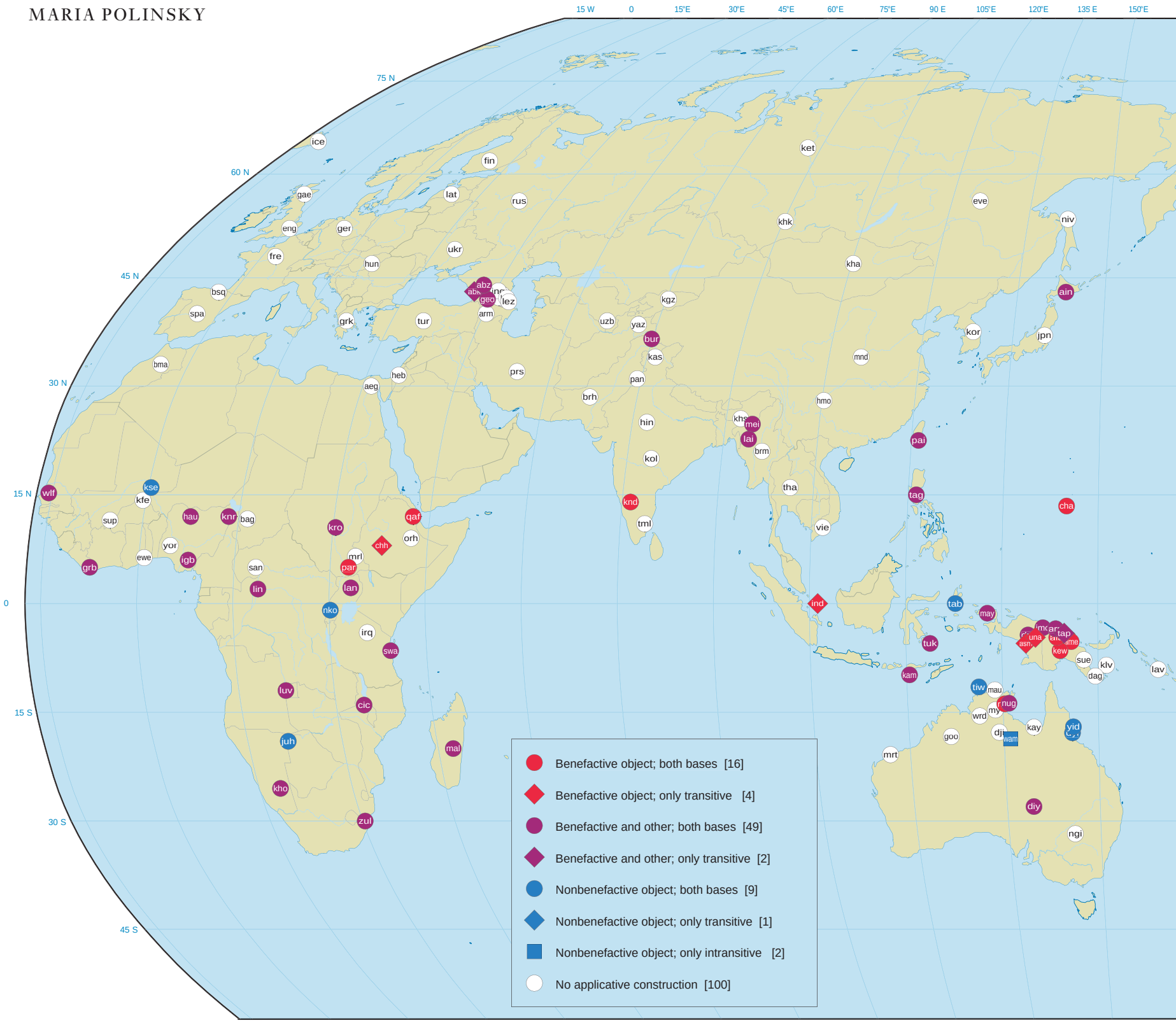
In some languages both objects, basic and applied, are accessible for passivization and relativization, can bind a reflexive, can trigger agreement on the verb, and/or can license coreferential deletion across clauses. Such languages are called "symmetrical". In other languages, only one object, either applied or basic, can show the relevant grammatical behaviours, while the other object is syntactically quite inert. Such languages are called "asymmetrical" (Woolford 1993; Alsina and Mchombo 1993). The asymmetrical type seems to be more common. The underlying structural superiority of the base object is typically claimed in the derivational account of applicatives (Baker 1988). In this analysis, applicative formation is the result of the incorporation of a preposition into a verb.

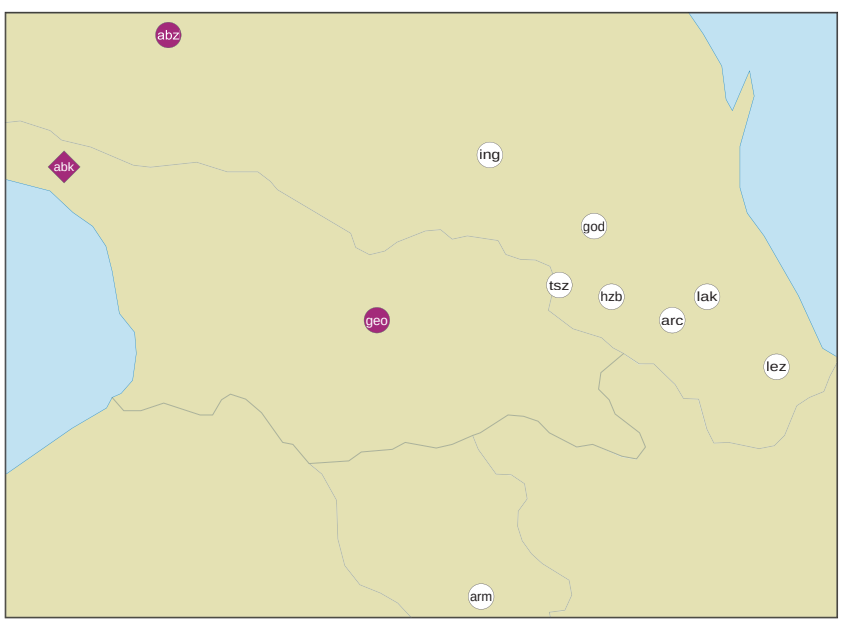
5.3 Iteration. Instances of multiple applicative marking and multiple applied objects are found in Koyraboro Senni, Bantu languages, Tukang Besi (and possibly other Austronesian languages), and Abaza. In Cahuilla, applicative formation is achieved either via prefixation (for adding a locative object) or via suffixation (for adding benefactive/recipient). In those cases where identical applicative morphemes allow iteration, as in Bantu, it is unclear whether their number is constrained by grammar or by processing limitations.

5.4 Semantics of the applied object. It is sometimes hard to tell from grammatical descriptions whether benefactive is differentiated from recipient/goal. It is also unclear how many languages merge benefactive and malefactive (the adversely affected object); impressionistically it seems more typical to keep them apart. In addition, the applied object can be mapped onto a possessor, whose place in the hierarchy of semantic roles is not quite clear. Benefactives, goals, and possessors are typically animate participants, which may create an impression that the applied object has to be an animate participant. Indeed, in some languages (e.g. Halkomelem) the referent of the applied object must be animate regardless of its semantic role (Gerdts 1988; 1993). In some languages, e.g. Kinyarwanda, the linear order of the base and the applied object varies depending on animacy (Kimenyi 1988; Polinsky and Kozinsky 1992; Polinsky 1995). Overall, animacy probably constitutes a separate dimension in the semantics of the applied object.

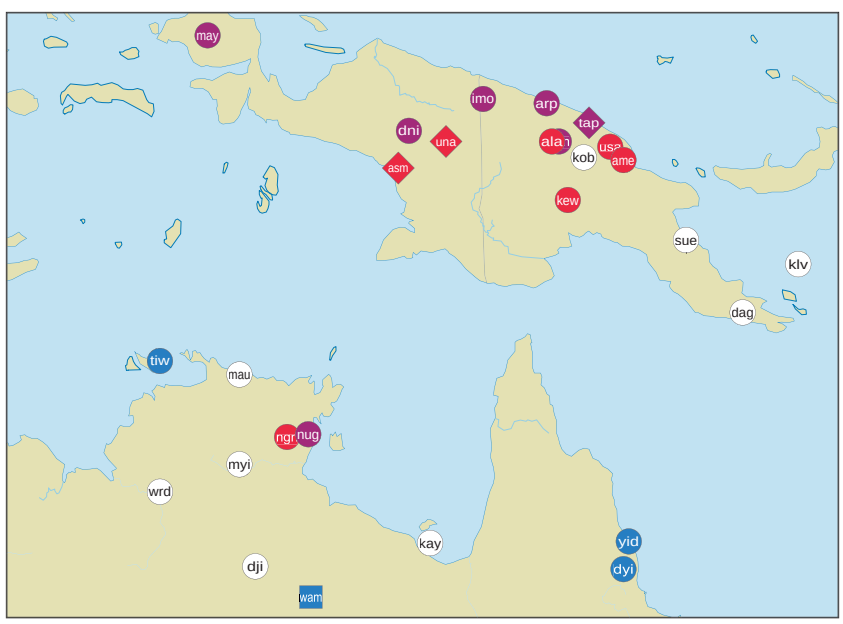
109 Applicative Constructions

MARIA POLINSKY

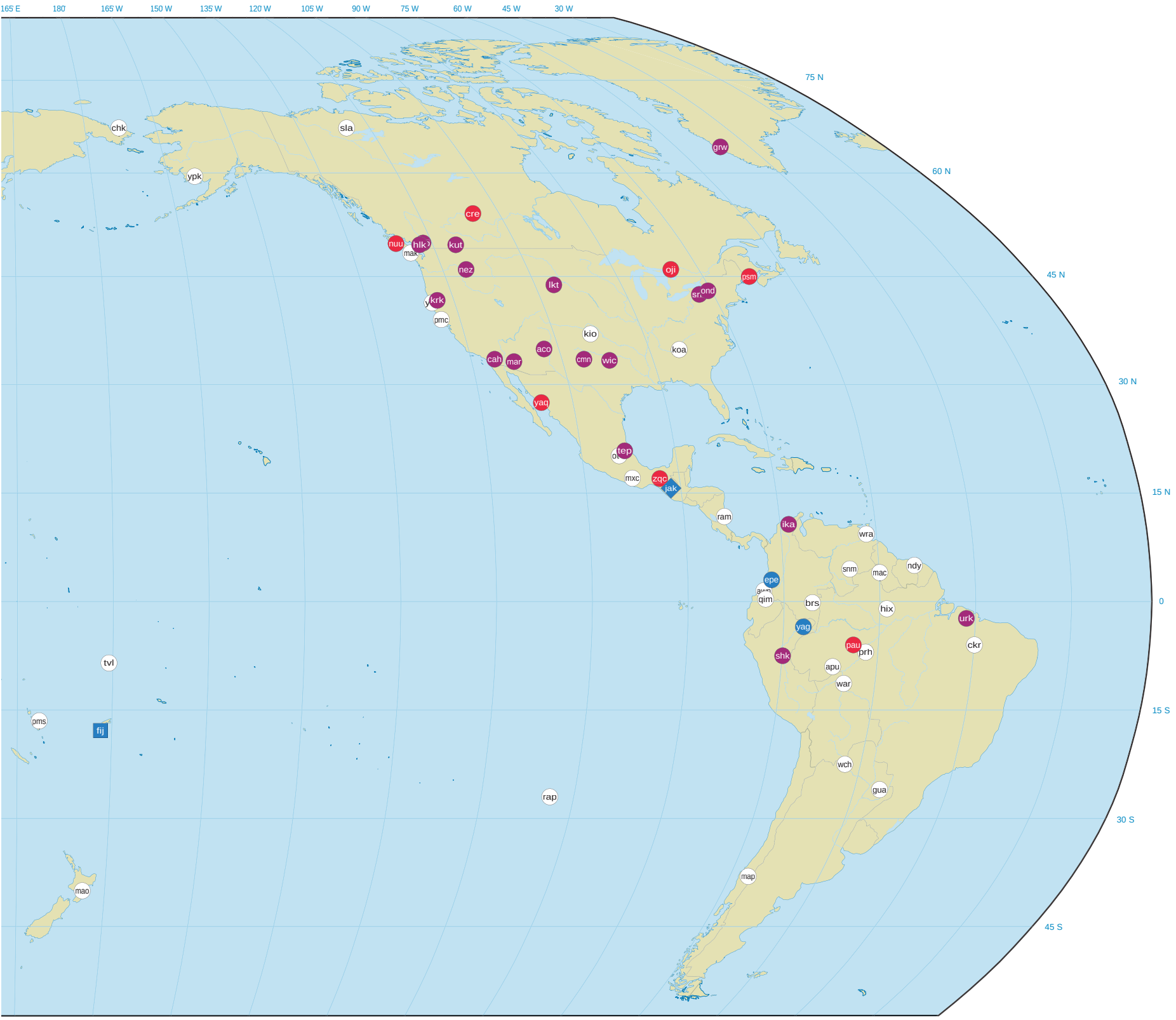


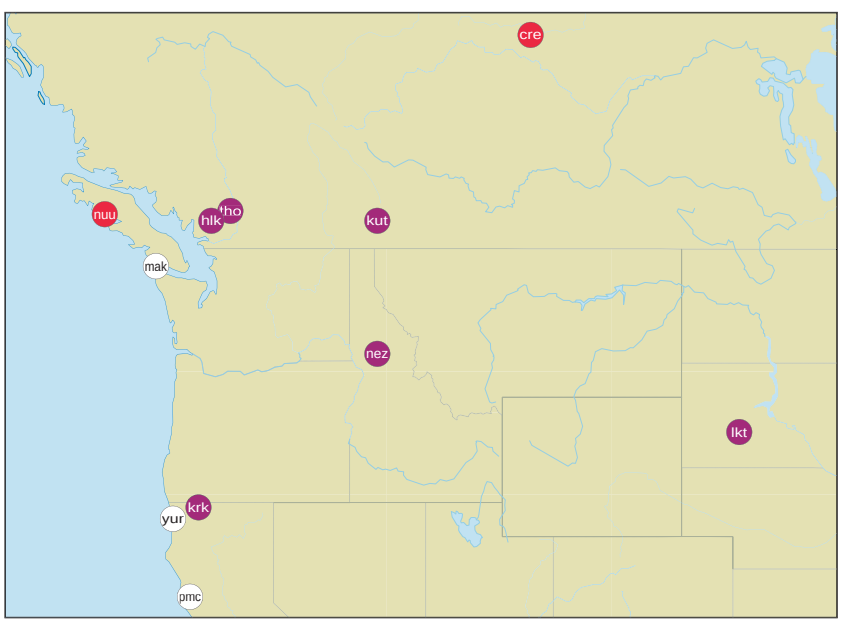


Caucasus

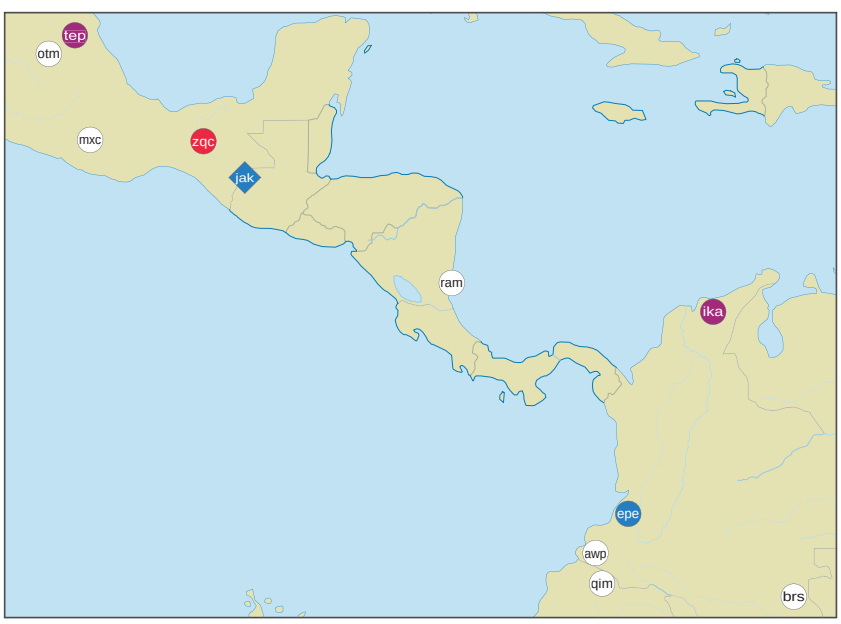


New Guinea and northern Australia





western North America



from southern Mexico to Columbia

110 Periphrastic Causative Constructions

JAE JUNG SONG

1 Defining periphrastic causative constructions

The causative construction is a linguistic expression which denotes a complex situation consisting of two component events (Comrie 1989: 165–6; Song 2001: 256–9): (i) the **causing event**, in which the **causer** does or initiates something; and (ii) the **caused event**, in which the **causee** carries out an action, or undergoes a change of condition or state as a result of the causer's action. The following sentence in Kinyarwanda (Bantu; Central East Africa) is such a linguistic expression.

(1) Kinyarwanda (Kimenyi 1980: 162)
umukoôbwa *y-a-tum-ye* *n-á-andik-a*
girl she-PST-cause-ASP I-PST-write-ASP
amábárúwa *meênshi*
letters many
'The girl caused me to write many letters.'

In (1), the causer (*the girl*) did something, and as a result of that action the causee (*me*) in turn carried out the action of writing many letters. Map 110 shows the geographical distribution of periphrastic causative constructions (also known as syntactic or analytic causative constructions in the literature).

Periphrastic causative constructions are causative expressions with the following three properties. First, the expression of the causer's action (or the **predicate of cause**) and the expression of the causee's action or change of condition or state (or the **predicate of effect**) must be in different clauses, as in (1). To put it differently, causative expressions, for the purposes of the map, must be **biclausal**. (The notion of biclausality being a continuum, however, the causative construction in one language (as in (5)) may be less biclausal than in another language (as in (1)); for further discussion, see Song 1996.) Second, the causer noun phrase and the predicate of cause (e.g. *umukoôbwa y-a-tum-ye* in (1)) must be "foregrounded", with the causee noun phrase and the predicate of effect (e.g. *n-á-andik-a* in (1)) "backgrounded". In (1), this is achieved by placing the clause expressing the causing event before the clause expressing the caused event. If one of the two clauses is embedded in the other, the foregrounding of the causer noun phrase and the predicate of cause is achieved by putting these two expressions in grammatically more "prominent" positions in the sentence than the causee noun phrase and the predicate of effect, as in (7) below. In (7), the causer noun phrase is the subject of the whole causative sentence, and the predicate of cause appears in the main clause of that sentence. In contrast, the causee noun phrase appears in a nonsubject position, and the predicate of effect is included in the nominalized (subordinate) clause. Third, the expression of the causer's action should be without specific meaning. In (1), for example, the predicate of cause *-tum-*, as opposed to the predicate of effect *-andik-*, lacks specific meaning; all that is expressed by the predicate of cause, *-tum-*, is the pure notion of causation.

Causative expressions such as (2), taken from Turkish, thus fall outside the purview of this map (but are the topic of Chapter 111).

(2) Turkish
Ali-Ø *Hasan-ı* *öl-dür-dü.*
Ali-SUBJ Hasan-DOBJ die-CAUS-PST
'Ali killed Hasan.'

In (2), the causer noun phrase, as opposed to the causee noun phrase, occupies the subject position; moreover, the expression of the causer's action—the causative suffix *-dür*—is without specific meaning. However, the question as to whether the predicate of cause is foregrounded in (2) does not arise at all, because there is only one predicate. In (2), the expression of the causer's action and the expression of the causee's change of state are both contained in one and the same predicate (or one and the same verb in this case). In other words, (2) is a **monoclausal** causative expression.

Biclausal causative expressions such as (3), taken from Spanish, are also not considered for the purposes of the present map.

(3) Spanish
Pedro *lo* *hizo* *porque* *Carmen* *vino*
Pedro it do.3SG.PST because Carmen came
a *la* *clase.*
to the class
'Pedro did it because Carmen came to the class.'

In (3), the causer noun phrase—if Carmen is regarded as the causer—is part of the subordinate clause, whereas the causee noun phrase—again, if Pedro is regarded as the causee—is part of the main clause. Moreover, the predicate of cause appears in the subordinate clause, while the predicate of effect appears in the main clause. To put it differently, the causer noun phrase and the predicate of cause are both placed in grammatically less prominent positions, with the causee noun phrase and the predicate of effect appearing in grammatically more prominent positions. Also note that the expression of the causer's action *vino* has a specific lexical content.

2 Defining the values

As indicated in the feature-value box, there are two types of periphrastic causative construction: the sequential type and the purposive type. The logically possible value of "no sequential type or purposive type" (i.e. lack of periphrastic causative constructions) is not listed in the feature-value box, because all languages seem to have some periphrastic means of expressing causation, and no languages have been satisfactorily described as completely lacking periphrastic causative constructions (see also below). Thus only three values are represented on the map:

1. Sequential type but no purposive type	35
2. Purposive type but no sequential type	68
3. Both sequential type and purposive type	15
total	118

As has already been noted, the periphrastic causative construction is biclausal in that the predicate of cause and the predicate of effect occur in different clauses. In the **sequential type**, the clause of cause and that of effect are juxtaposed strictly in that order, with or without a linking element between them. In other words, the first clause denotes the causing event and the second clause the caused event (hence the term sequential).

The Kinyarwanda causative in (1) is an example of the sequential type, without use of a linking element. Kobon (Madang, Trans-New Guinea; Papua New Guinea) is another good example of this type.

(4) Kobon (Davies 1981: 165)
nipe *g-aj-ɨp* *yad* *mab* *rɨb-pin*
3SG do-DUR-PST.3SG 1SG tree cut-PFV.1SG
'He made me cut the tree.'

The sequential type with use of a linking element is exemplified by Amele (Madang; Papua New Guinea). Note that the linking element used in (5) is a so-called switch reference affix *-ce-* (see Song 1996: 35–49 for further exemplification of the sequential type).

(5) Amele (Roberts 1987: 222)
ija *od-ude-ce-min* *na* *qete-i-a*
1SG do-3SG-SR-1SG tree cut-3SG-TOD.PST
'I made him cut the tree.'

The **purposive type** also involves two clauses, one representing an event carried out for the purpose of realizing another event denoted by the other clause (hence the term purposive). The sense of purpose or goal present in this subtype can be signalled by (i) verbal markings such as future tense, irrealis, subjunctive mood, incompletive aspect, etc.; (ii) dative, allative, or purposive case markers; or (iii) purposive particles.

Swahili (Niger-Congo) is an example of the purposive type. In the causative sentence in (6), the predicate of effect, which must appear in subjunctive mood, stands in contrast with the predicate of cause in the past tense. In other words, the subjunctive mood marking is utilized here to engender the sense of purpose or goal.

(6) Swahili (Vitale 1981: 153)
Ahmed *a-li-m-fanya* *mbwa* *a-l-e*
Ahmed he-PST-him-make dog he-eat-SBJV
samaki *mkubwa*
fish large
'Ahmed made the dog eat a large fish.'

In other languages of the purposive type, case markers such as dative, allative, purposive, and the like may be utilized to signal the sense of purpose or goal. Basque is such a language.

(7) Basque (Saltarelli et al. 1988: 221)
Mikel-ek *Jon* *Edurne-ri* *liburu-a*
Mikel-ERG Jon(ABS) Edurne-DAT book-SG.ABS
eros-te-ra *behar-tu* *z-u-en*
buy-NMLZ-ALL make-PFV 3SG.ERG-AUX-PST
'Mikel made Jon buy a book for Edurne.'

Note that in (7) the nominalized (subordinate) clause of effect as a whole is "flagged" by the allative case marker *-ra* (cf. *parke-ra* 'to the park'). The purposive type of causative construction in Modern Greek relies on a subjunctive particle (i.e. *na*) to signal the sense of purpose or goal, as in (8) (see Song 1996: 49–67 for further exemplification of the purposive type).

(8) Modern Greek (Joseph and Philippaki-Warburton 1987: 171)
ékan-a *ton* *jáni* *na* *fíj-i*
made-1SG.ACTIVE the John.ACC SBJV leave-3SG
'I made John leave.'

When reading Map 110 in conjunction with Map 111, readers will notice that many of the languages appearing on the latter are not represented on the former. As mentioned also in the companion chapter, this is largely because the primary sources for these languages discuss only the nonperiphrastic causative construction without even indicating whether or not a periphrastic causative construction is in use as well. (This explains why Map 111 displays more languages than Map 110 does.) Moreover, it is well known that the morphological process involved in the nonperiphrastic causative is rarely completely productive (e.g. Nedjalkov and Silnitsky 1973; Song 1996: 170–2). In view of this, it is very likely that most of the languages shown as employing only the nonperiphrastic causative may also have other means of expressing causation, i.e. the periphrastic causative construction. These points are worth bearing in mind when interpreting the two maps together. Still, a good number of languages are identified on the maps as making use of either the sequential type or the purposive type (or even both) in addition to the nonperiphrastic causative.

3 Geographical distribution

As noted in §2, there are gaps in the database concerning use of the periphrastic causative construction. Thus it may seem somewhat premature to make comments on geographical patterns of the sequential and purposive types. Nonetheless, a few tentative observations can be made.

In the eastern part of Africa stretching from Ethiopia to South Africa (including Afro-Asiatic, Nilo-Saharan, and Niger-Congo languages), the purposive type is much more common than the sequential type. In the western part of Africa (again including Afro-Asiatic, Nilo-Saharan, and Niger-Congo languages), on the other hand, the sequential type seems to be the dominant type, although there are exceptions, namely Koyraboro Senni, Supyire, Vai, and Yoruba. Three of these exceptions also make use of the sequential type, however.

In a vast area covering Europe, the Caucasus, and the Middle East, and extending almost into Central Asia, the purposive type predominates over the sequential type. The dominance of the purposive type in this area may possibly link with the predominance seen in the eastern part of Africa.

In India and South-East Asia taken together, the purposive type is the predominant type in the west, with the sequential type being more common in the east. What remains to be determined—once more data become available—is whether the western half of this combined area may ultimately meet the vast area alluded to earlier, that is, Europe, the Caucasus, the Middle East, and the eastern part of Africa, all of which display a strong preference for the purposive type over the sequential type.

In Australia, the purposive type is clearly the dominant type, the only exponent of the sequential type being Diyari, spoken in South Australia. In New Guinea, on the other hand, the sequential type seems to prevail over the purposive type, although there are many gaps in the data. The sequential type in New Guinea also seems to extend eastwards and westwards, although Oceania has its own share of the purposive type.

There are too many gaps in the data to be able to make suggestions about the geographical distribution of the sequential and purposive types in North and South America, although one may venture to say that there is a concentration of purposive-type languages from the bottom of North America to the top of South America.

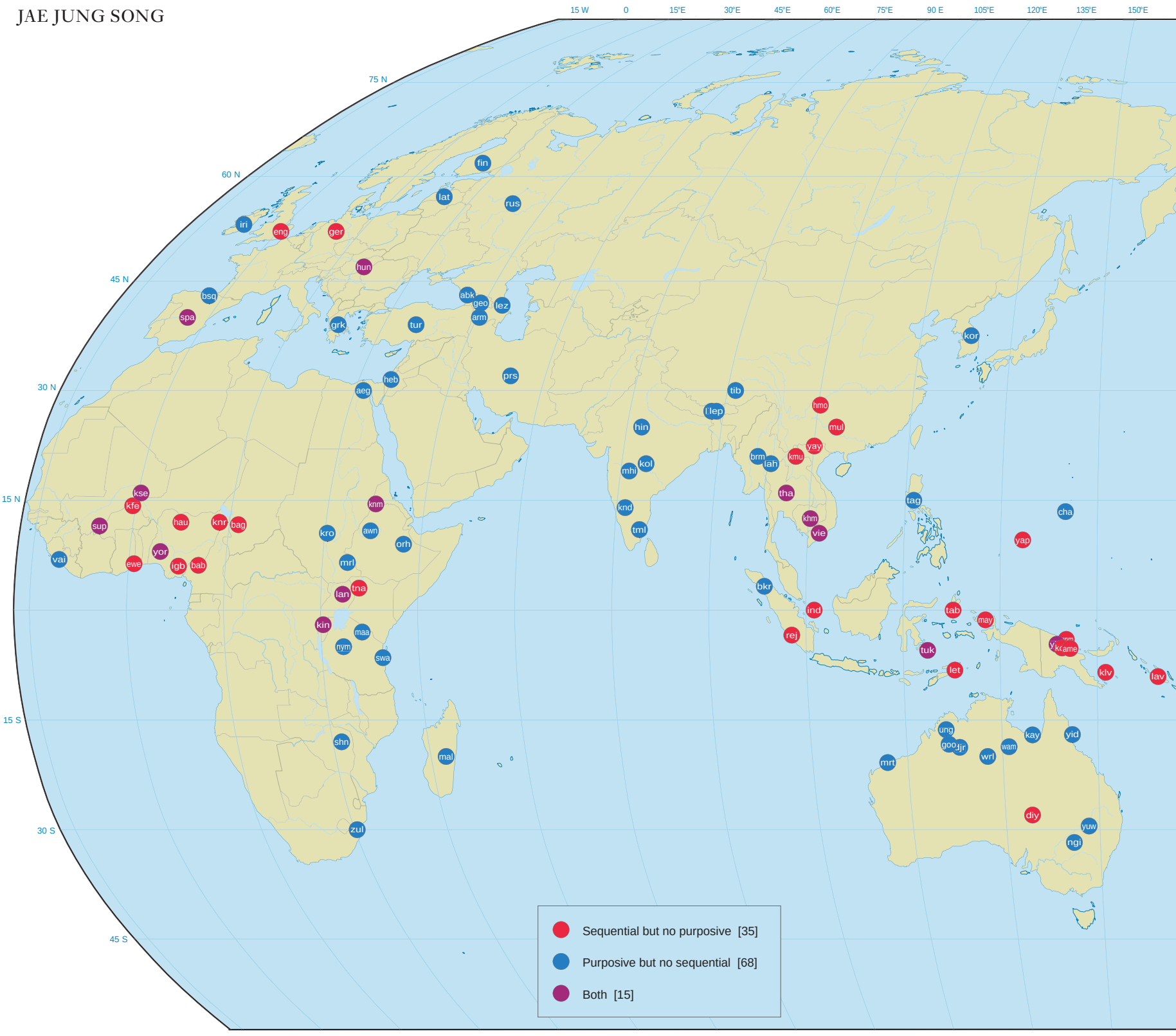
There are a small number of languages with both the sequential type and the purposive type. Indo-China (Khmer, Thai, and Vietnamese) and the western part of Africa (Koyraboro Senni, Supyire, and Yoruba) may contain potential clusters of such languages, but more data are required to confirm these clusters.

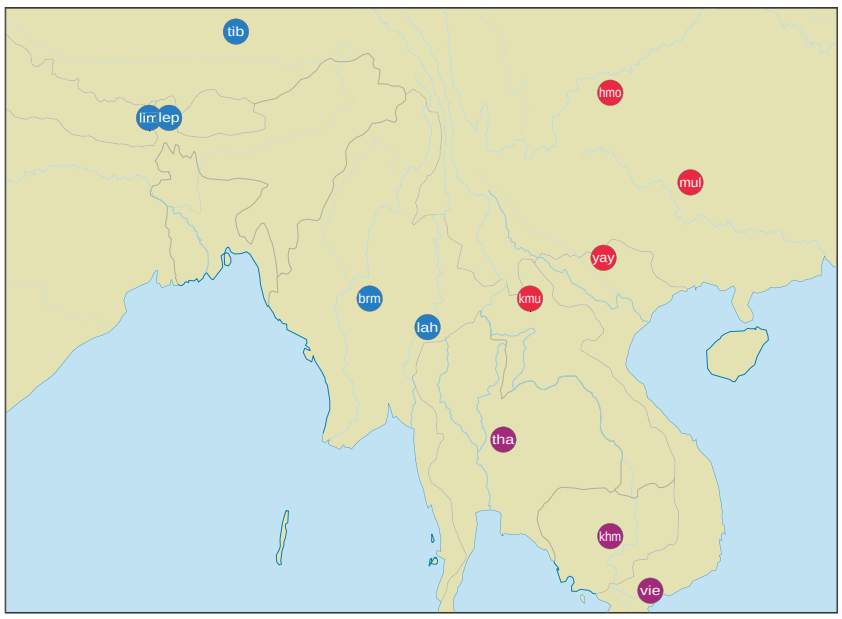
4 Theoretical issues

For the past three decades or so, the causative construction has been a recurrent research topic in linguistics. Most research has focused on the nonperiphrastic causative construction with respect to the grammatical relation of the causee noun phrase, and on the causative as a valency or grammatical relation-changing operation (Comrie 1976b and Dixon 2000; also see Song 1996: ch. 6). Thus it is not inaccurate to say that previous theories of causative constructions have dealt mainly with the nonperiphrastic causative construction, as a result of which the variety of the periphrastic causative construction has been understated. Song (1996) can be regarded as an attempt to redress the balance, as it were, by highlighting and documenting the rich cross-linguistic variation of the periphrastic causative construction.

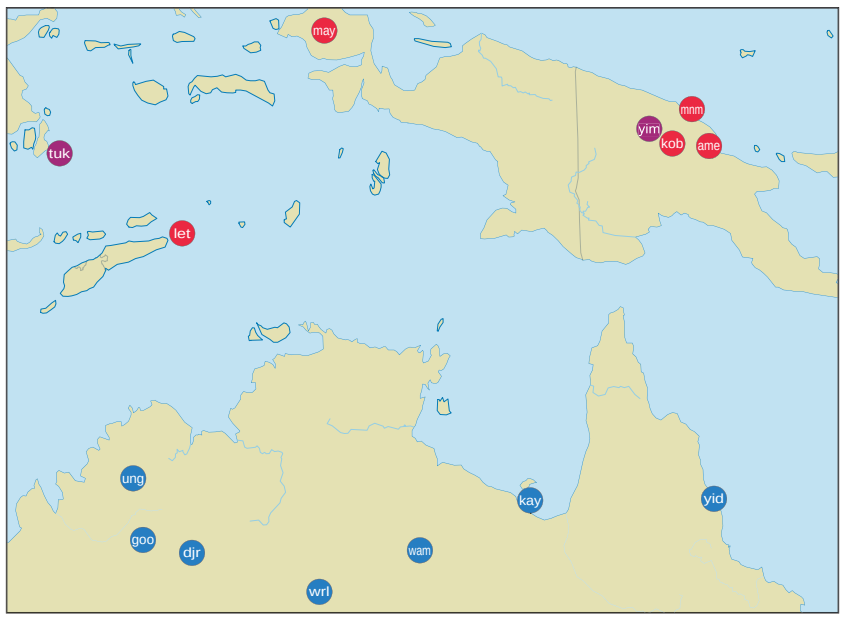
110 Periphrastic Causative Constructions

JAE JUNG SONG

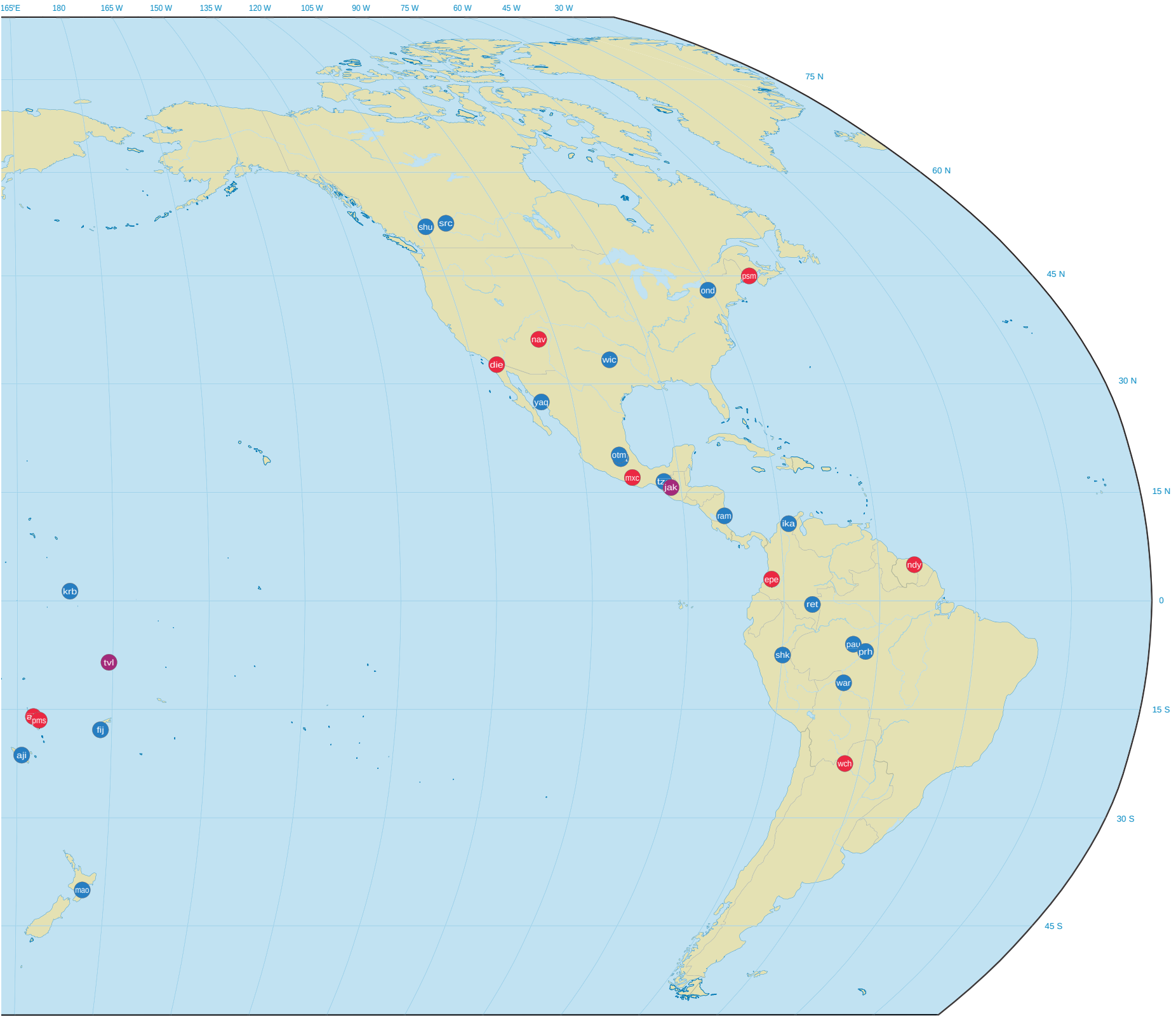


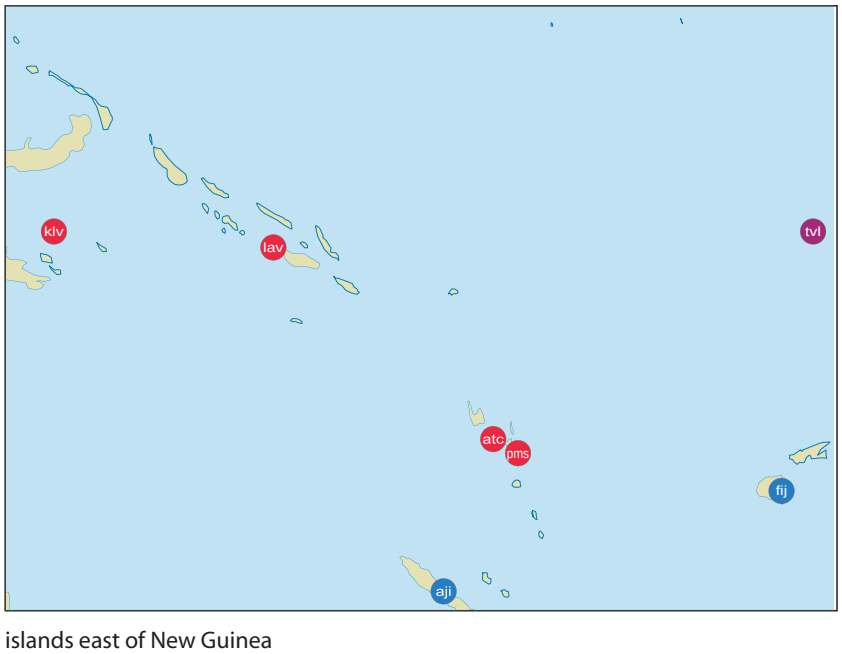


from eastern Nepal to Vietnam

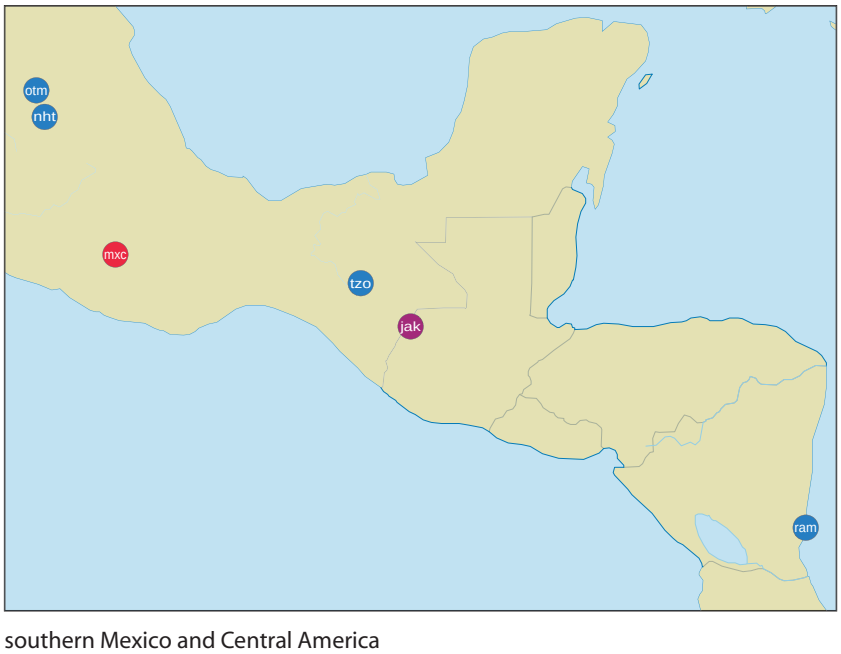


New Guinea and northern Australia





islands east of New Guinea



southern Mexico and Central America

111 Nonperiphrastic Causative Constructions

JAE JUNG SONG

1 Defining nonperiphrastic causative constructions

The causative construction is a linguistic expression which denotes a complex situation consisting of two component events (Comrie 1989: 165–6; Song 2001: 256–9): (i) the **causing event**, in which the **causer** does or initiates something; and (ii) the **caused event**, in which the **causee** carries out an action, or undergoes a change of condition or state as a result of the causer's action. The following Japanese sentence is such a linguistic expression.

(1) Japanese
Kanako ga Ziroo o ik-ase-ta
Kanako NOM Ziro ACC go-CAUS-PST
'Kanako made Ziro go.'

In (1), the causer (*Kanako*) did something, and as a result of that action the causee (*Ziro*) in turn carried out the action of going. Map 111 shows the geographical distribution of nonperiphrastic causative constructions.

Nonperiphrastic causative constructions are causative expressions with the following three properties. First, the expression of the causer's action (e.g. *-ase* in (1)) and the expression of effect (e.g. *ik-* in (1)) must both be contained in one and the same predicate, which may consist of one or more verbs (e.g. one verb *ik-ase-* in (1), or two verbs *me ŋò* in (6) below). To put it differently, such causative expressions must be **monoclausal**. Second, the causer noun phrase must occupy a grammatically more "prominent" position (e.g. the subject in (1)) than the causee noun phrase (the object in (1)). Third, the expression of the causer's action, be it an affix or a separate verb, should be without specific meaning. In (1), the expression of cause *-ase*, as opposed to the expression of effect *ik-*, lacks specific meaning; all that is expressed by *-ase* is the pure notion of causation.

Causative expressions such as (2), taken from Manam (Oceanic; Papua New Guinea), are not considered here, because the causer's action and the effect are expressed by different predicates. These predicates in turn appear in different clauses. (Causative expressions such as (2) are the topic of Chapter 110.)

(2) Manam (Lichtenberk 1983a: 449)
wása ʔúsi i-emaʔ-í-be i-moaʔúsu
wind cloth 3SG.REAL-cause-3SG.OBJ-and 3SG.REAL-move
'The wind made the loincloth move.'

Monoclausal causative expressions such as (3), taken from Tuvaluan (Polynesian), are also not included in the present discussion, because the causer's action, if expressed at all, is not contained in the predicate (although the notion of effect is).

(3) Tuvaluan (Besnier 2000: 325)
te paala e tasi ne sao
the kingfish NPST one PST escape
ia Aaifoou
because.of Aifou
'The kingfish got away because of Aifou.'

2 Defining the values

As indicated in the feature-value box, there are two types of nonperiphrastic causative construction: the morphological type and the compound type. Thus four values are represented on the map:

○	1. No morphological type or compound type	23
●	2. Morphological type but no compound type	254
●	3. Compound type but no morphological type	9
●	4. Both morphological type and compound type	24
	total	310

The **morphological type** involves a morphological process which applies directly to a basic verb (Song 1996: 21–6; Dixon 2000: 33–4). Typically, this morphological process consists in affixation: in the morphological type the predicate—or the causative verb—is made up of a basic verb and a causative affix (Comrie 1989: 167–8). In Japanese, for example, the causative suffix *-(s)ase* applies to basic verbs to derive causative verbs, as illustrated in (1) above. Compare (1) with (4), the noncausative counterpart of (1).

(4) Japanese
Ziroo ga ik-u
Ziro NOM go-PRES
'Ziro goes.'

The causative element can take the form not only of suffixes (e.g. *-(s)ase* in Japanese), but also prefixes (e.g. *r-* in Abkhaz), infixes (e.g. *-y-* in Lepcha), and circumfixes (e.g. *a- . . . -ineb* in Georgian) (Song 1996: 21–8).

The morphological process may also involve an internal change in vowel or consonant quality (e.g. Lahu *dɔ̀* 'drink' versus *tɔ* 'give (i.e. cause) to drink'), gemination of a consonant in the basic verb (e.g. Egyptian Arabic *mawat* 'die' versus *mawwit* 'kill'), internal vowel lengthening (e.g. Kashmiri *marun* 'die' versus *mārun* 'kill'), reduplication of the basic verb (e.g. Korana *xa* 'learn' versus *xa-xa* 'teach'), or even a tonal change (e.g. Lahu *câ* 'eat' versus *cā* 'feed'). There is one other variation on the morphological type worth mentioning here. In Russian, for example, it is noncausative, and not causative, verbs that bear an additional morpheme, i.e. the anticausative suffix *-sja* (e.g. *lomat'* 'to break [something]' versus *lomat'-sja* '[for something] to break'). For the purposes of Map 111, the Russian causative verb will also be regarded as an instantiation of the morphological type: in comparison with the noncausative verb, it can be thought to contain a zero causative morpheme. In the sample, there are only a small number of languages which display variations on the morphological type like those found in Lahu, Egyptian Arabic, Kashmiri, Korana, and Russian.

In the **compound type**, the causer's action is expressed by a separate verb instead of a morphological element, but that verb must appear next to a basic verb so that other elements are not able to intervene between the two. Thus the two verbs behave as a single predicate (i.e. a compound causative verb), very much as the basic verb and the causative suffix in (1) are a single predicate. Romance languages such as French and Spanish are described as languages having the compound type. In French, for example, the two verbs in question, albeit separate lexical items, must appear next to each other, so that the causee noun phrase (and the object noun phrase of the basic verb) is (are) prevented from breaking up the unit, as illustrated in (5). The negative element *pas* or adverbials can in fact intervene between the two verbs, but this does not detract from the fact that normally nothing comes in between them. (For further discussion, see Comrie 1976b: 296–303.)

(5) French
Je le lui ferai lire.
I it.ACC her.DAT make.FUT read
'I'll make her read it.'

In Eastern Kayah Li (Tibeto-Burman; Myanmar) the verb of cause and that of effect must abut on each other, as in (6). Note that the causer noun phrase *ʔa* and the causee noun phrase *phúcè* appear before and after the sequence of both verbs.

(6) Eastern Kayah Li (Solnit 1997: 69)
*ʔa **me** ŋò phúcè*
he do laugh child
'He made the child laugh.'

It is not merely the physical adjacency of the verbs of cause and effect that allows one to recognize the compound type, but also the ability of these two verbs to form a grammatical unit. This is demonstrated, for example, by the difference between Tamil (Dravidian; southern India and northern Sri Lanka) and Kobon (Madang; Papua New Guinea). The Tamil causative sentence in (7) is not regarded as being of the compound type in spite of the fact that the verb of cause and that of effect are adjacent to each other.

(7) Tamil (Asher 1985: 155)

naan	*avane*	*veekamaa*	*oo*ɻ*-a*	*vacceen*
I	he.ACC	quickly	run–PURP	cause.PST.1SG

'I made him run quickly.'

The reason why (7) is not taken as an example of the compound type is that the verb of effect, *oo*ɻ*-a*, actually contains purposive marking, *-a*. The presence of this purposive marking in the verb of effect indicates that the two verbs do not form a unit, and that *oo*ɻ*-a* is part of a subordinate clause of purpose. Thus (7) should instead be regarded as an example of the periphrastic causative construction. In fact, the two verbs in (7) end up next to each other mainly because of the verb-final word order of Tamil. The verb of effect, being in the final position of the subordinate clause, appears right before the verb of cause (i.e. the verb of the main clause). Kobon, also a verb-final language, has the verb of cause and that of effect next to each other, but unlike in Tamil, there is evidence that the two verbs form a kind of grammatical unit, which is characteristic of the compound type. Consider:

(8) Kobon (Davies 1981: 164)

a. *mab* *dudu.g-öp*
tree be.bent–PERF.3SG
'The tree is bent.'

b. *yad* *mab* *dudu.g-i* ***yu-bin***
1SG tree be.bent throw–PERF.1SG
'I bent the tree.'

Example (8a) is a noncausative sentence, whereas (8b) is a corresponding causative sentence. (Note that in (8b) the verb of cause *yu-* has lost its original meaning of throwing; recall that the expression of the causer's action should be without specific meaning.) In (8b), the causer's action is expressed by a separate verb *yu-*, which appears immediately after the verb of effect *dudu.g-i*. In (8b), the single aspect marker on the verb of cause has its scope over the whole unit, supporting the view that the two verbs form a compound causative verb.

Languages such as Modern Greek, Kilivila, Krongo, and Maybrat use neither the morphological type nor the compound type. Thus readers may wonder how the causative situation, as defined at the beginning of §1, is expressed in these languages. They depend on periphrastic causative constructions. Languages like these are a minority, however.

When reading Map 111 in conjunction with Map 110, readers will notice that many of the languages appearing on the former are not represented on the latter. As mentioned also in the companion chapter, this is largely because the primary sources for many of these languages discuss only the morphological type without even indicating whether or not a periphrastic causative construction is in use as well. (This explains why Map 111 displays more languages than Map 110 does.) Moreover, it is well known that the morphological process involved in the nonperiphrastic causative is rarely completely productive (e.g. Nedjalkov and Silnitsky 1973; Song 1996: 170–2). In view of this, it is very likely that most of the languages shown as employing only the morphological type may also have other means of expressing causation, i.e. the periphrastic causative. These points must be borne in mind when interpreting the two maps together. Still, a good number of languages are identified on the maps as making use of both the morphological type and the periphrastic causative.

3 Geographical distribution

In view of the prevalence of the morphological type in the world's languages, it is not inappropriate to describe its geographical patterning in negative terms, that is to say, with respect to its absence. There are languages lacking the morphological type in all major geographical areas: Africa, Eurasia, South-East Asia–Oceania, Australia–New Guinea, North America, and South America. Within Eurasia, languages without the morphological type are found only in Europe, namely Modern Greek and Irish. In each of the major geographical areas, however, languages without the morphological type are too few or scattered to reveal clear geographical patterns, except for possible clusters of such languages in South-East Asia (Mandarin, Hmong Njua, Eastern Kayah Li, Mulao, Thai, Vietnamese, and Yay) and northwestern Australia (Djaru, Gooniyandi, and Ungarinjin).

There are a very small number of languages with the compound type and without the morphological type in the sample. In all areas where these languages are located, however, they are too few and far between to suggest any distinct geographical patterns.

As pointed out in §2, there are languages without either the morphological type or the compound type, for example, Modern Greek, Kilivila, Krongo, and Maybrat. Given the very small number of languages with the compound type and without the morphological type, it comes as no surprise that the geographical patterns suggested above for languages without the morphological type also more or less hold true for languages without the morphological type and the compound type. Thus there are two potential clusters of languages lacking the nonperiphrastic causative: South-East Asia (Hmong Njua, Mulao, Thai, and Yay) and northwestern Australia (Djaru, Gooniyandi, and Ungarinjin).

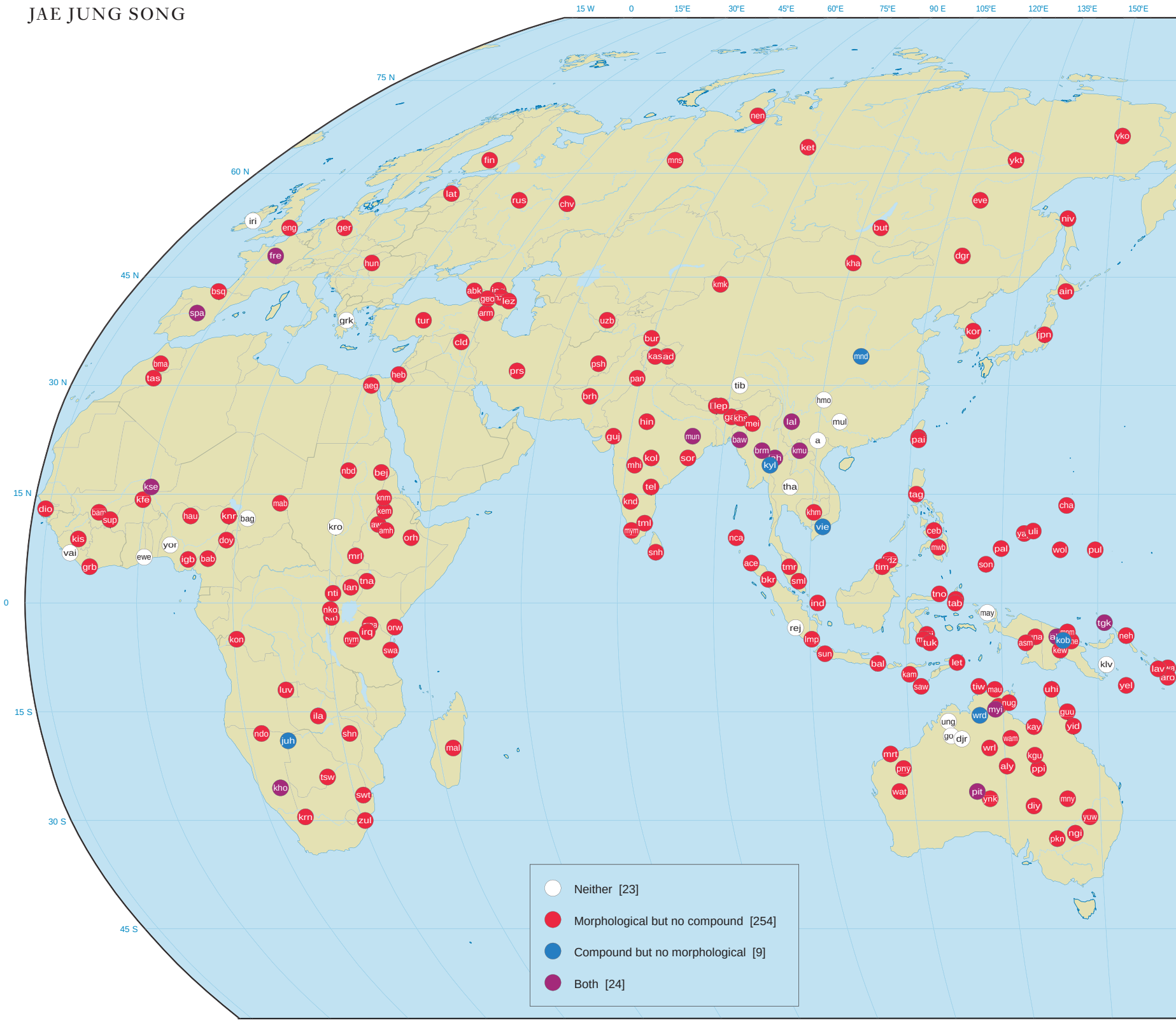
Finally, languages with both the morphological type and the compound type are not numerous, but there are at least two areas where such languages seem to cluster together: South-East Asia (Bawm, Burmese, Khmu', Lahu, and Lalo) and the southeastern part of North America (Biloxi, Koasati, and Tunica). However, more data will be required to substantiate these clusters.

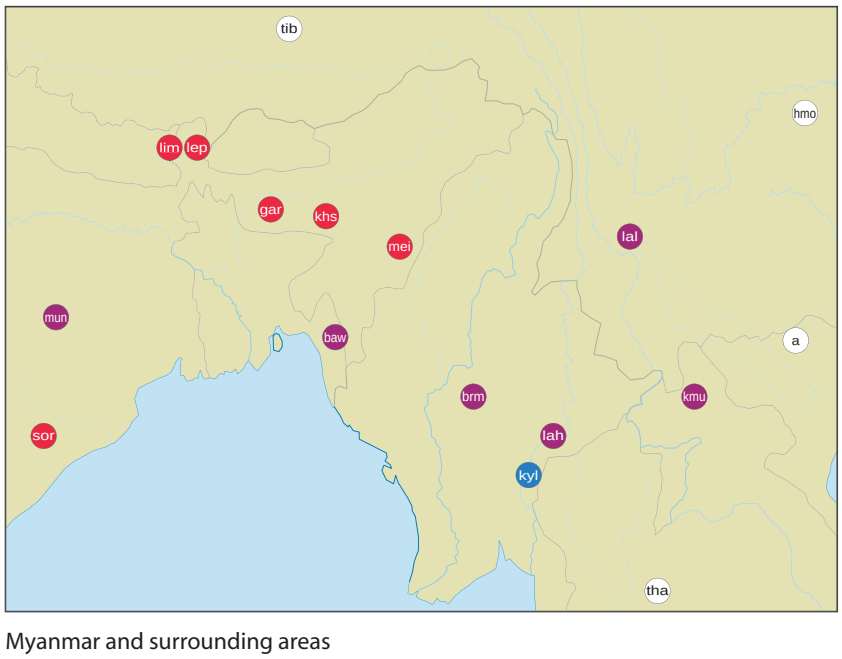
4 Theoretical issues

For the past three decades or so, the causative construction has been a recurrent research topic in linguistics. Most research has focused on the morphological type (and also the compound type as a variation on the morphological type) with respect to the grammatical relation of the causee noun phrase, and on the causative as a valency or grammatical relation-changing operation (Comrie 1976b and Dixon 2000; also see Song 1996: ch. 6). It is claimed that the grammatical relation of the causee noun phrase in the morphological type of causative—which is assumed to consist of a matrix clause and an embedded clause at some abstract level—can be predicted by appealing to a hierarchy of grammatical relations, although some linguists have argued that the morphological type of causative is modelled conceptually and grammatically on simple clauses. Formal theories also have proposed different approaches to representation of causativization as a grammatical relation-changing operation. These approaches can be characterized broadly as lexical or syntactic.

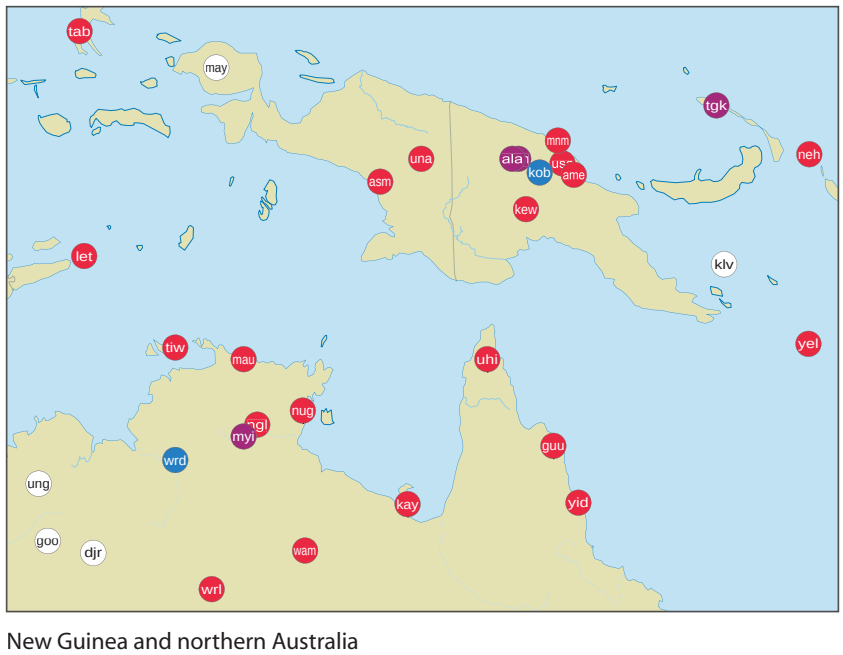
111 Nonperiphrastic Causative Constructions

JAE JUNG SONG

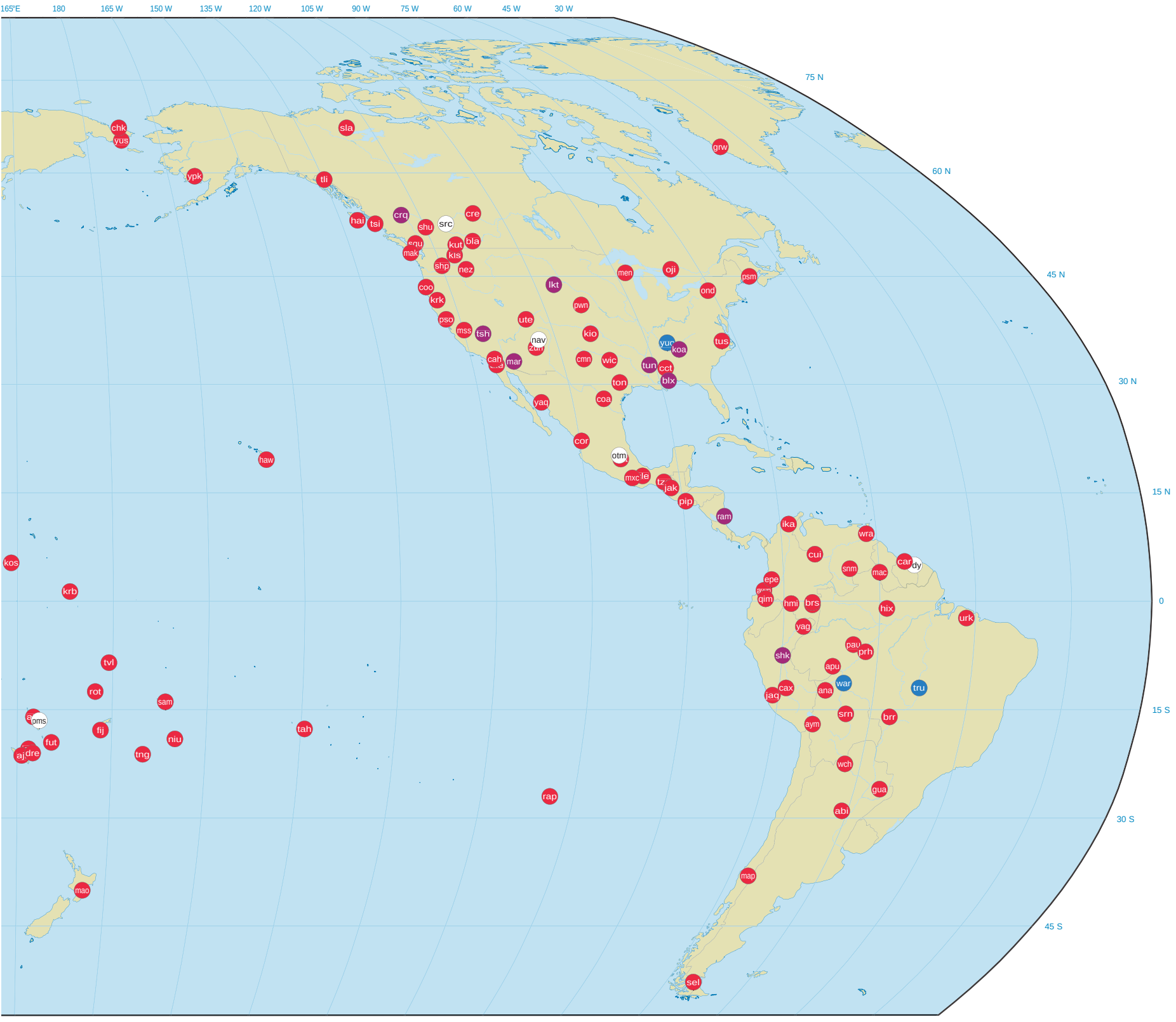


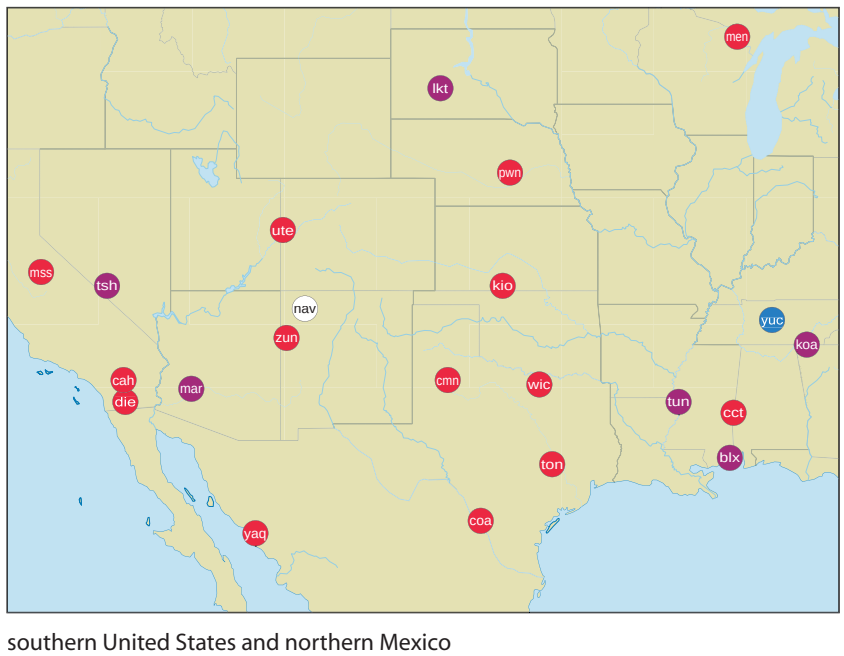


Myanmar and surrounding areas

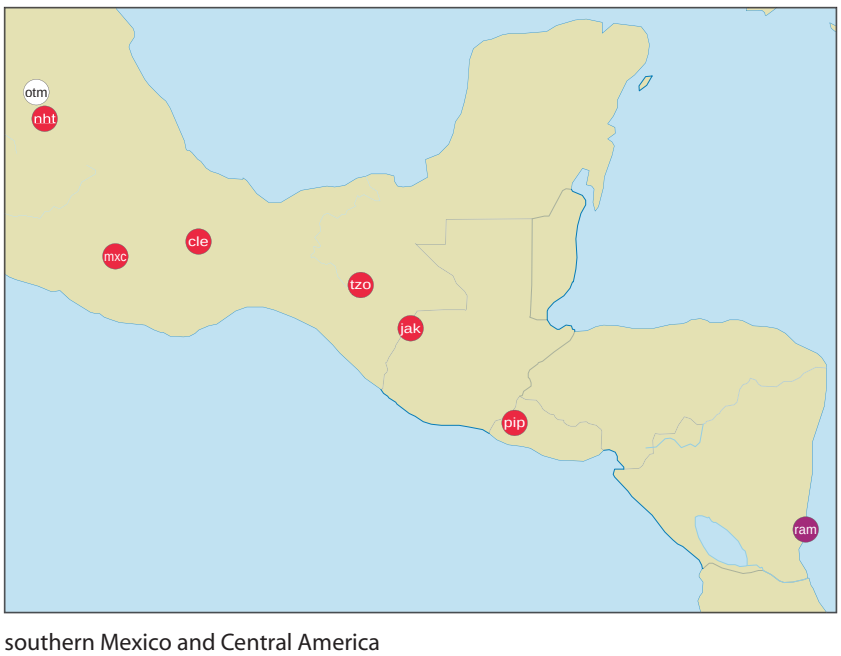


New Guinea and northern Australia





southern United States and northern Mexico



southern Mexico and Central America

112 Negative Morphemes

MATTHEW S. DRYER

1 Defining the values

This map shows the nature of morphemes signalling clausal negation in declarative sentences. By clausal negation is meant the simple negation of an entire clause; (1b) is the negative clause corresponding to the affirmative clause in (1a):

(1) a. *John is eating pizza.*
b. *John is* ***not*** *eating pizza.*

Not considered here are noun phrase negation (*No students were present*), negative pronouns (*Nobody came*), or negative adverbs (*She never eats pizza*) (though see Map 115 on negative pronouns). The map is based on the expression of negation in declarative sentences; thus, where a language employs a distinct construction in imperative sentences, that construction is ignored here (see Map 71).

1.	Negative affix	339
2.	Negative particle	477
3.	Negative auxiliary verb	45
4.	Negative word, unclear if verb or particle	65
5.	Variation between negative word and affix	19
6.	Double negation	66
	total	1011

All of the ways of indicating negation involve negative morphemes. This contrasts, for example, with a variety of ways of signalling a polar question (Map 116) that do not involve interrogative morphemes, such as changes in word order, intonation, and the complete absence of any signal that a sentence is a question. There are no known instances of languages in which negation is realized by a change in word order or by intonation, and all languages have negative morphemes.

The first type shown on the map consists of languages that express negation by means of a **negative affix** attached to the verb, as in (2) from Kolyma Yukaghir (Siberia, Russia):

(2) Kolyma Yukaghir (Maslova 2003: 492)
met *numö-ge* ***el-****jaqa-te-je*
1SG house-LOC NEG-achieve-FUT-INTR.1SG
'I will not reach the house.'

The next three types all involve languages which express negation by separate words. The first type of negative word is a **negative particle**, like the English word *not* or the word *pày* in (3) from Musgu (Chadic, Afro-Asiatic; Cameroon).

(3) Musgu (Meyer-Bahlburg 1972: 186)
à *sədà* *cécébè* ***pày***
3SG.M know jackal NEG
'He didn't see the jackal.'

The second type of negative word is a word that inflects as a verb, and which can be considered a type of **auxiliary verb**, since it normally must accompany another verb. In Finnish, for example, the negative word is an auxiliary verb, and inflects for the person and number of the subject, as in (4), while what is semantically the main verb occurs in a nonfinite participle form.

(4) Finnish (Sulkala and Karjalainen 1992: 115)
e-n *syö-nyt* *omena-a*
NEG-1SG eat-PTCP apple-PART
'I didn't eat an apple.'

In Grebo (Kru, Niger-Congo; Liberia), the negative word's status as a verb is reflected by the fact that it takes the tense inflection that occurs on the verb in the corresponding affirmative sentence. For example, in the affirmative sentence in (5a), the verb *du*^{1} 'pound' occurs with a past-before-yesterday tense suffix *-da*^{2}. In the corresponding negative sentence in (5b), this tense suffix occurs on the negative word *yi*^{21} instead.

(5) Grebo (Innes 1966: 55, 106)
a. *ne*^{1} *du*^{1}*-da*^{2} *bla*^{4}
1SG.SUBJ pound-PST.BEFORE.YEST rice
'I pounded rice before yesterday.'
b. *ne*^{1} ***yi***^{21}***-da***^{2} *bla*^{4} *du*^{1}
1SG.SUBJ NEG-PST.BEFORE.YEST rice pound
mɔle
Monday
'I did not pound rice on Monday.'

For many languages, it is easy to decide whether a negative word is a particle or a verb: if the negative word does not occur with inflectional morphology associated with verbs in the language, it is treated here as a particle, while if it does occur with inflectional morphology associated with verbs in the language, then it is treated here as a negative auxiliary verb. In some languages, however, it is more difficult to decide whether the negative word should be considered a verb or not, if verbs occur with little or no inflectional morphology or if the inflectional morphology that does occur on verbs may not be semantically appropriate for a negative word even if that negative word is a verb. Such languages are treated here as instances of the fourth type shown on the map, as languages with **negative words, but where it is unclear if the negative is a verb or a particle**. For example, in Maori (Polynesian; New Zealand), the negative is uninflected, but so is the verb, as illustrated in (6).

(6) Maori (Bauer 1993: 140)
kaahore *taatou* *e* *haere* *ana* *aapoopoo*
NEG 1PL.INCL T/A move T/A tomorrow
'We are not going tomorrow.'

Because of the absence of inflectional morphology on verbs in Maori, it is difficult to decide on the basis of superficial evidence whether the negative word should be considered a verb or not. In some languages, there may be good syntactic reasons for classifying a negative word as a verb. For example, in Boumaa Fijian (Oceanic), the negative word has the syntactic properties of a main verb, and the accompanying predication is placed in a complement clause that functions as the subject of the negative verb, as in (7).

(7) Boumaa Fijian (Dixon 1988: 40)
e ***sega*** *[ni* *la'o* *o* *Jone]*
3SG NEG COMP go ART John
'John is not going.'
(literally: 'That John is going is not the case.')

However, in many languages with little or no inflectional morphology on verbs, it is not obvious whether or not the negative word should be considered a verb even by syntactic criteria; morever, the criteria assumed for this map are purely morphological. Thus even if there seems to be good syntactic (but not morphological) reason to say that the negative word in a language is a verb, it is classified as a language with a negative word where it is unclear if the negative word is a verb or particle.

The fifth type consists of languages which have more than one negative construction, one in which the negative is a separate word, and one in which it is an affix. For example, Rama (Chibchan; Nicaragua) has two negative constructions, one with a preverbal negative particle, as in (8a), the other with a negative suffix on the verb, as in (8b).

(8) Rama (Grinevald n.d.: 183, 185)

a. *nkiikna-lut uut* ***aa*** *kain-i*
man-PL dory NEG make-TNS
'The men don't make a dory.'

b. *i-sik-**taama***
3-arrive-NEG
'He did not arrive.'

Languages like Rama are coded as languages with **variation between a negative word and a negative affix**.

The last type shown on the map is languages with **double negation**, where the negative construction involves two simultaneous morphemes, one preceding the verb or verb stem, the other following. In Standard French, for example, the negative construction involves both a preverbal negative particle *ne* and a postverbal negative particle *pas*, as in (9).

(9) French
Je ***ne*** *vois* ***pas*** *la lune.*
1SG NEG see.1SG NEG the moon
'I do not see the moon.'

In some languages with double negation, the negative construction involves a prefix and a suffix on the verb, as in (10) from Izi (Igboid, Niger-Congo; Nigeria).

(10) Izi (Meier et al. 1975: 218)
ó ***tó**-òmé-**dú*** *ré*
3SG NEG-do-NEG well
'He does not do well.'

In still other languages, one of the negative morphemes is an affix on the verb while the other is a separate word, as in (11), from Ma (Adamawa-Ubangi, Niger-Congo; Democratic Republic of Congo).

(11) Ma (Tucker and Bryan 1966: 130)
tá**-mù-sùbù-li *nɔ́ŋgbɔ́* ***nyɔ̀***
NEG-1SG-eat-PST meat NEG.1SG
'I did not eat meat.'

No attempt is made on the map to distinguish the different types of negative morphemes that occur in languages with double negation, i.e. whether the negative morphemes involved are negative affixes or negative words or a combination of these.

Some languages employ double negation only under certain circumstances. For example, Mupun (Chadic; Nigeria) employs an obligatory clause-final negative word as well as an optional clause-initial negative word, as illustrated in (12), where the parentheses around the initial word indicate its optionality.

(12) Mupun (Frajzyngier 1993: 353)
(ba) *kə̀ n=se lua nyer* ***kas***
NEG PERF 1SG=eat meat bird NEG
'I did not eat the bird meat.'

Languages like Mupun with optional double negation are not shown on the map as involving double negation but rather according to the properties of the obligatory negative mopheme. Similarly, French is coded on the map as employing a negative particle rather than as having double negation, since the preverbal particle is optional in colloquial French.

In some languages, the morphemes that code negation also code other grammatical features of the clause as well. For instance, the example in (13) from Rumu (Turama-Kikorian, Trans-New Guinea; Papua New Guinea) illustrates a verbal suffix that codes both negation and future tense.

(13) Rumu (Petterson 1999: 24)
*hope eitariki hërö koi-**moi***
bamboo straight properly stand.upright-NEG.FUT
'Bamboo will not stand up straight.'

Negative particles also occasionally vary for tense-aspect, as in the examples in (14) from Puluwat (Oceanic; Micronesia); contrast the affirmative perfective marker *ya* in (14a) with the negative perfective marker *há* in (14b).

(14) Puluwat (Elbert 1974: 85)

a. *yi ya kiililó*
1SG.SUBJ PFV.AFF hungry
'I'm hungry.'

b. *yi* ***há*** *kúleey*
1SG.SUBJ PFV.NEG know
'I don't know.'

In Hausa (Chadic), the negative particles vary not only for tense and aspect but also for pronominal features of the subject. For example, the first word *ban* in the main clause in (15) codes negation, completive aspect, and first singular subject, while the word *bài* in the relative clause codes negation, completive aspect, and third singular masculine subject.

(15) Hausa (Newman 2000: 359)
bàn *ga yārò-n dà*
NEG.COMP.1SG see boy-DEF REL
bài *tàimàki Lādì* ***ba***
NEG.COMP.3SG.M help Ladi NEG
'I didn't see the boy who didn't help Ladi.'

One might think that negative words like those in Puluwat and Hausa are reminiscent of negative auxiliary verbs, as illustrated in (4) and (5) above. However, in (4) and (5), the inflections on the negative word are inflections which otherwise occur on verbs in the language, so that they provide a basis for saying that the negative word is a verb. In (14) and (15), in contrast, verbs do not in general inflect for tense-aspect or for subject; hence the negative words are best viewed as a set of nonverbal words that vary for more than one grammatical feature of the clause, one of which is negation.

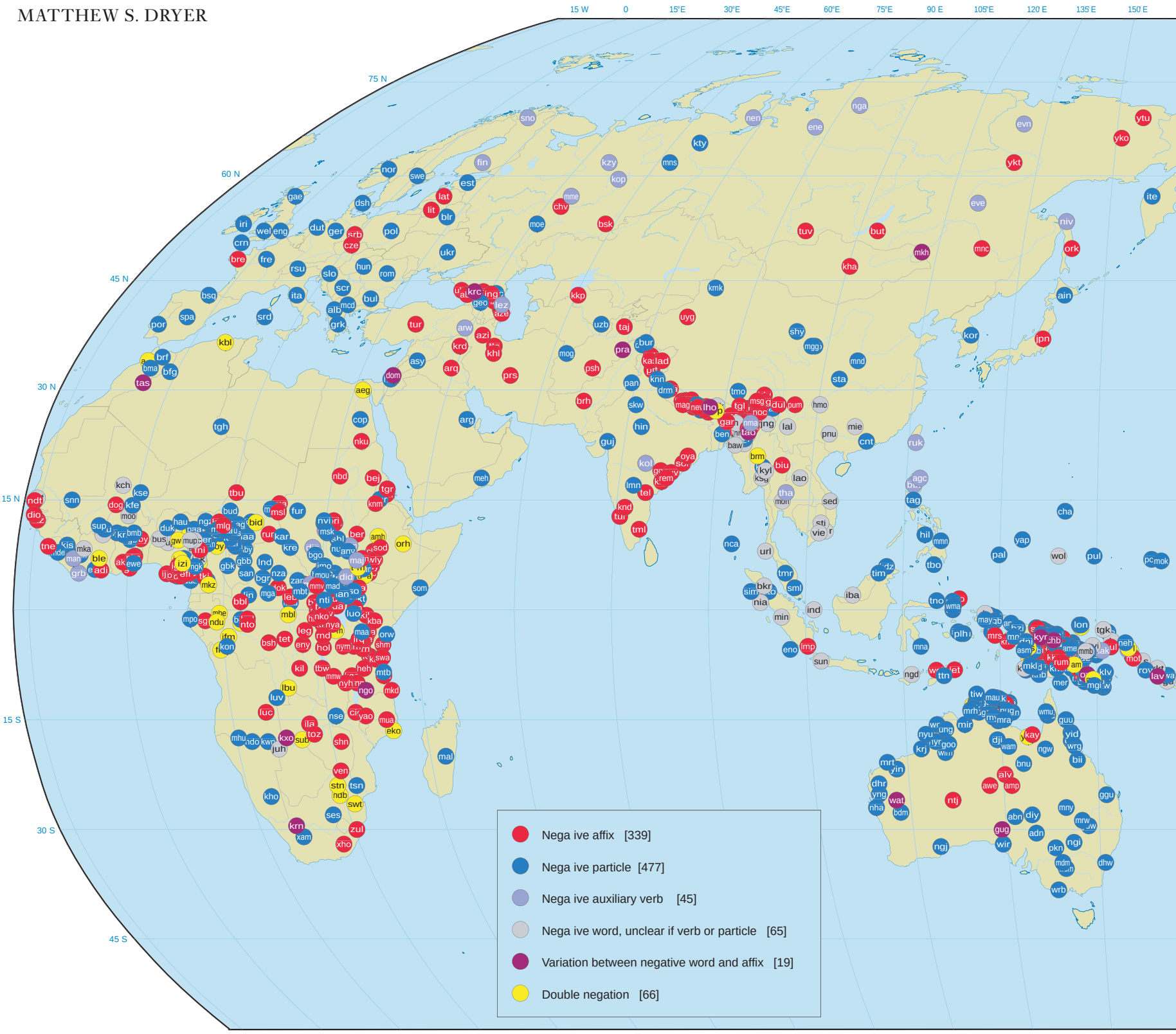
There are many other dimensions to the typology of negation than those covered by this map. See Maps 113 and 114 for additional aspects of the typology of negation.

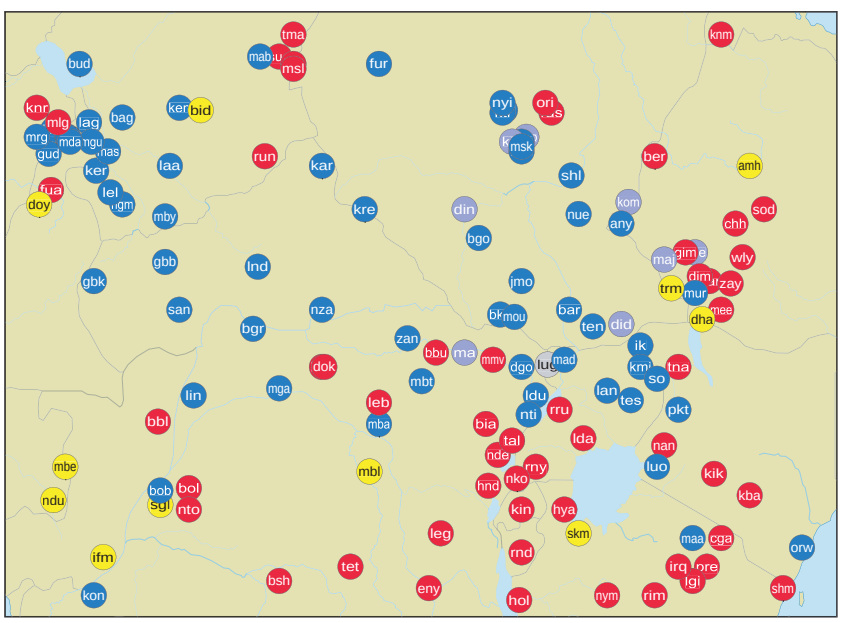
2 Geographical distribution

The geographical patterns are somewhat less striking here than on some other maps. Both negative particles and negative affixes are widely distributed throughout the world. Languages with negative auxiliaries are less common, but show a striking frequency across northern Eurasia, stretching from Finland to western Siberia. Languages with other negative words are common in South-East Asia, but this largely reflects the isolating nature of languages in this area, making it difficult to determine whether negative words are verbs or not. Double negation is also widespread, except for Europe and much of Asia, but seems particularly frequent in Africa.

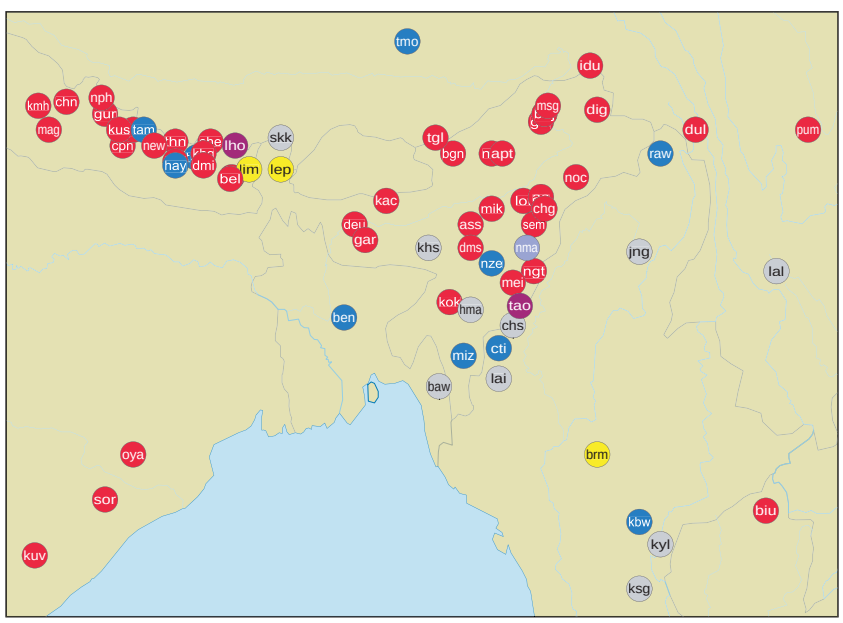
112 Negative Morphemes

MATTHEW S. DRYER

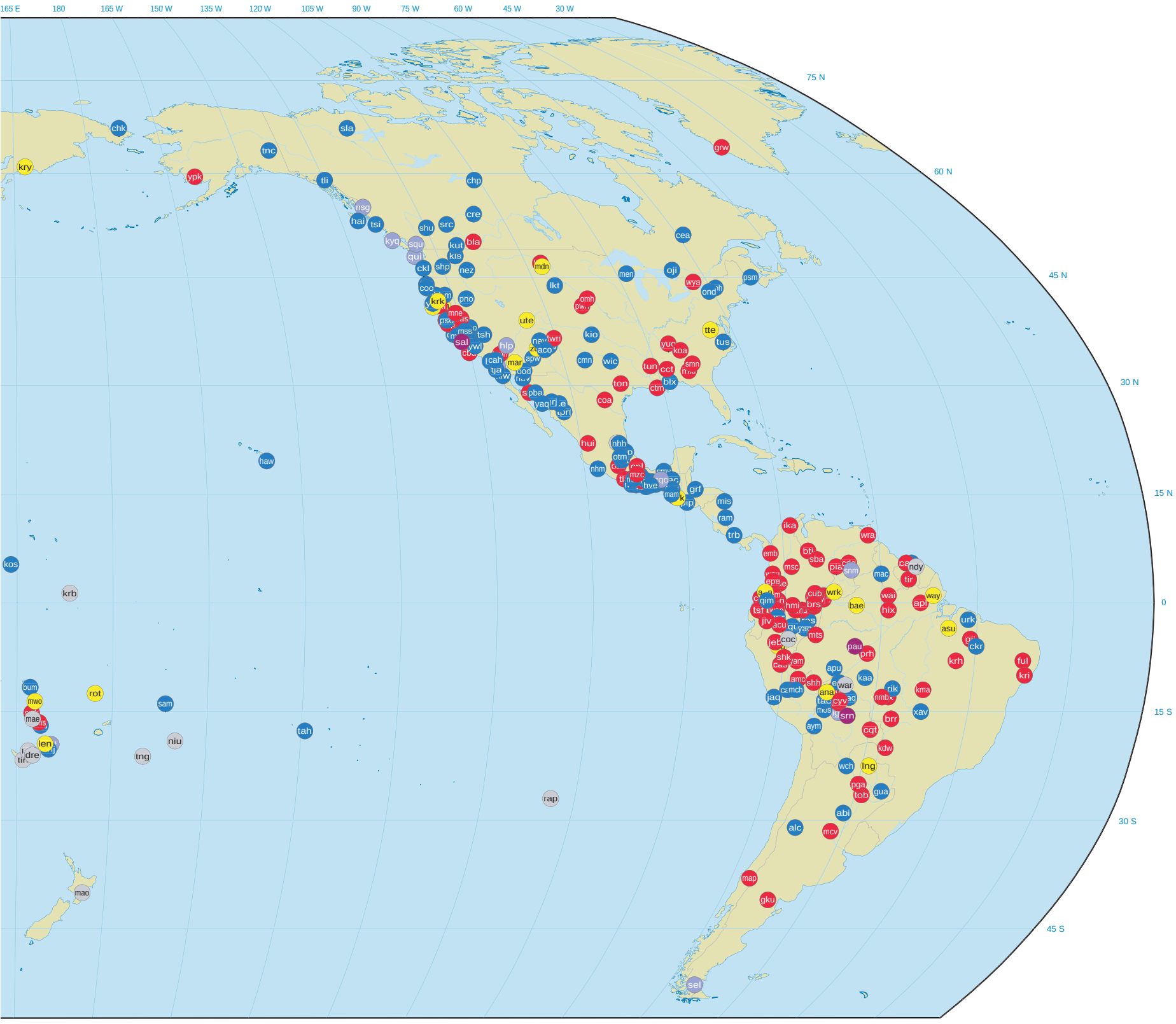


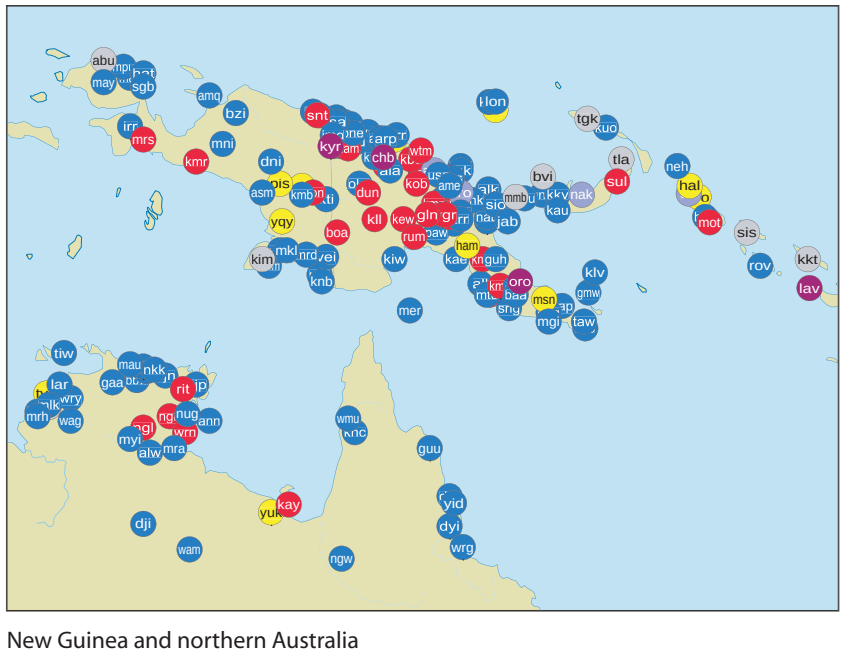


central Africa

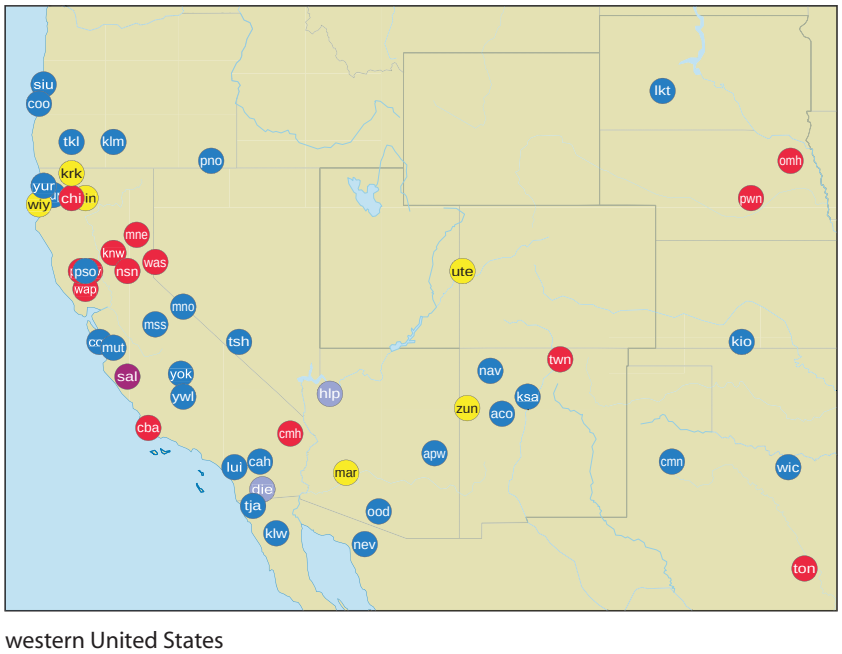


Nepal, eastern India, and adjacent areas





New Guinea and northern Australia



western United States

113 Symmetric and Asymmetric Standard Negation

MATTI MIESTAMO

1 Definition of values

Standard negation can be defined as the basic way (or ways) a language has for negating declarative verbal main clauses. Negative constructions that fall outside standard negation include the negation of existential, copular, or nonverbal clauses, the negation of subordinate clauses, and the negation of nondeclarative clauses like imperatives (see Chapter 71). These negatives are not taken into account here, but it is of course possible that languages use their standard negation constructions for the negation of these clause types too.

This map shows how symmetric and asymmetric standard negation are distributed among the languages of the world. In **symmetric negation** the structure of the negative is identical to the structure of the affirmative, except for the presence of the negative marker(s). In **asymmetric negation** the structure of the negative differs from the structure of the affirmative in various other ways too, i.e. there is asymmetry between affirmation and negation. Affirmative and negative structures can be symmetric or asymmetric in two ways: there can be (a)symmetry either between the affirmative and negative **constructions**, or between the **paradigms** that the affirmative and negative constructions form. Symmetric negative constructions do not differ from the corresponding affirmative constructions in any other way than by the presence of the negative marker(s), whereas asymmetric negative constructions show structural differences in comparison to the corresponding affirmative constructions. In symmetric paradigms, all (verbal) categories or forms have corresponding affirmative and negative forms, whereas in asymmetric paradigms such one-to-one correspondences do not obtain.

Three types of languages are distinguished on the map: Type Sym: languages where standard negation is always symmetric; Type Asy: languages where standard negation is always asymmetric; and Type SymAsy: languages where both symmetric and asymmetric standard negation is found.

○	1. Symmetric standard negation only: Type Sym	114
●	2. Asymmetric standard negation only: Type Asy	53
○	3. Symmetric and asymmetric standard negation: Type SymAsy	130
	total	297

The structural differences, i.e. the asymmetry, can show up in different grammatical domains, and according to the domain and nature of the asymmetry, subtypes of asymmetric negation can be established. In the present map the subtypes of asymmetric negation play no role; these subtypes will be discussed in Chapter 114. I will now give examples of the Types Sym, Asy, and SymAsy. Examples of asymmetric negatives can also be seen in Chapter 114.

In **Type Sym** negation is always symmetric, i.e. no asymmetry is found—neither constructional nor paradigmatic. In German (see 1a–b) the negative construction is symmetric since only the negative marker *nicht* is added to the corresponding affirmative and no other structural differences are found between the affirmative and the negative.

(1) German (personal knowledge)

a. *ich singe*
I sing.1SG
'I sing.'

b. *ich singe nicht*
I sing.1SG NEG
'I do not sing.'

c. *singen* 'to sing', 1SG

	AFFIRMATIVE	NEGATIVE
PRES	*ich singe*	*ich singe nicht*
PST	*ich sang*	*ich sang nicht*
PERF	*ich habe gesungen*	*ich habe nicht gesungen*
PLUPERF	*ich hatte gesungen*	*ich hatte nicht gesungen*

The small sample of the German verbal paradigm given in (1c) illustrates the fact that all affirmative verbal forms have a corresponding negative form, i.e. all forms can be negated, and the negative paradigm is thus symmetric with the affirmative one. No asymmetry, constructional or paradigmatic, is found, and German is thus classified as a language of Type Sym.

Standard negation is also expressed with symmetric negative constructions in Taba (Austronesian; Halmahera, Indonesia; see (2)) and in Shipibo-Konibo (Panoan; Peru; see (3)).

(2) Taba (Bowden 2001: 335)

a. *n=han ak-la*
3SG=go ALL-sea
'She's going seawards.'

b. *n=han ak-la te*
3SG=go ALL-sea NEG
'She's not going seawards.'

(3) Shipibo-Konibo (Pilar Valenzuela, p. c.)

a. *rono-ra kako-nko ka-ke*
Rono.ABS-EVD Caco-ALL go-COMP
'Rono went to Caco.'

b. *rono-ra kako-nko ka-yama-ke*
Rono.ABS-EVD Caco-ALL go-NEG-COMP
'Rono did not go to Caco.'

In Taba the negative particle *te* is simply added to the positive clause and there are no further changes in the structure. In Shipibo-Konibo the negative suffix *-yama* appears on the verb with no further changes. These symmetric negative constructions are used in all standard negation contexts, and all affirmatives can be negated with these constructions in these languages. No constructional or paradigmatic asymmetry is thus found in Taba or Shipibo-Konibo, and they belong to Type Sym.

In **Type Asy** standard negation is always asymmetric. In Finnish (see (4)) negation is marked by a construction where the negative verb *e-* carries the marking of person-number, and the lexical verb loses its finiteness, appearing in the connegative form in the present (see (4a–b)) and in the past participle in the past (see (4c–d)). There are thus structural differences between the negative and the affirmative in addition to the presence of the negative marker; the negative construction is therefore analysed as asymmetric.

(4) Finnish (personal knowledge)

a. *tule-n*
come-1SG
'I am coming.'

b. *e-n tule*
NEG-1SG come.CONNEG
'I am not coming.'

c. *tul-i-n*
come-PST-1SG
'I came.'

d. *e-n tul-lut*
NEG-1SG come-PST.PTCP
'I did not come.'

The paradigm is symmetric in Finnish since all affirmatives can be negated with the asymmetric negative verb construction. But as all negatives use the asymmetric construction, negation is always asymmetric, and Finnish is a language of Type Asy.

The Burmese standard negation construction involves a combined prefix and suffix *ma- -bû* (see (5)). This negative suffix *-bû* replaces the tense-aspect-mood markers used in the affirmative.

(5) Burmese (Cornyn 1944: 12–13)

a. *θwâ-dé*
go-ACT
'((S)he) goes, went.'

b. *θwâ-mé*
go-POT
'((S)he) will go.'

c. *θwâ-bí*
go-PERF
'((S)he) has gone.'

d. *ma-θwâ-bû*
NEG-go-NEG
'((S)he) does/did/will not go, has not gone.'

As the negative markers are not simply added to the corresponding affirmative, the construction is asymmetric. In Burmese the paradigm is also asymmetric: the affirmative paradigm makes a distinction between actual, potential (future), and perfect, but the distinction is lost in the negative since the negative construction replaces the suffixes marking these categories. No symmetric constructions being found in Burmese, it is a language of Type Asy.

For a language to qualify as belonging to Type Asy, it is sufficient that every negative construction in the language be asymmetric. The paradigm can of course be symmetric or more or less asymmetric, but as the negative construction is always asymmetric, there are no instances of symmetric standard negation.

In languages of **Type SymAsy** both symmetric and asymmetric standard negation are found. Standard negation is sometimes symmetric, i.e. at least some negative constructions are symmetric, but there is also some asymmetry to be found, either in some negative constructions or somewhere in the paradigm. This is an intersection of the two types of standard negation within the same language. In Lezgian (Nakh-Daghestanian; eastern Caucasus; see (6)), the negative suffix *-č* appears in standard negation.

(6) Lezgian (Haspelmath 1993: 127, 245)

a. *xürünwi-jr-i* *ada-waj* *meslät-ar* *qaču-zwa*
villager-PL-ERG he-ADEL advice-PL take-IMPF
'The villagers take advice from him.'

b. *xürünwi-jr-i* *ada-waj* *meslät-ar* *qaču-zwa-č*
villager-PL-ERG he-ADEL advice-PL take-IMPF-NEG
'The villagers do not take advice from him.'

c. *fi-zwa-j*
go-IMPF-PST
'was going'

d. *fi-zwa-č-ir*
go-IMPF-NEG-PST
'was not going'

The negative construction is symmetric in non-past imperfectives (6a, b). In past imperfective verb forms (6c, d), past tense is marked differently in the affirmative and the negative, and the negative construction is thus asymmetric. Lezgian belongs to Type SymAsy.

In Maung (Yiwaidjan; Northern Territory, Australia) the negative clause is obligatorily marked for the irrealis category (see (7d)), whereas the affirmative can make a distinction between realis and irrealis (see (7a–c)).

(7) Maung (Capell and Hinch 1970: 67)

a. *ŋi-udba-ŋ*
1SG.3-put-R.PST.PUNCT
'I put [it]'

b. *ŋi-udba-ŋuŋ*
1SG.3-put-R.PST.CONT
'I was putting [it]'

c. *da* *ŋi-udba-nji*
if 1SG.3-put-IRR.PST
'if I put [it]'

d. *marig* *ŋi-udba-nji*
NEG 1SG.3-put-IRR.PST
'I did not put [it]'

The negative construction simply adds the negative marker *marig* to the non-negative irrealis, and the construction is thus symmetric. There is paradigmatic asymmetry since the realis–irrealis distinction is lost in the negative. There are symmetric negatives in Maung, but some asymmetry is also found, and the language thus belongs to Type SymAsy. Note that the punctual–continuous distinction made in the affirmative (7a, b) is also lost in the negative, but this distinction is lost in the irrealis in general, and the loss is thus not directly due to negation.

As seen above, a language can have different standard negation constructions used in different contexts (e.g. with different tense-aspect categories) or in free variation. Furthermore, the negative paradigm can be symmetric or asymmetric, and this is partly independent of the type of negative construction(s) the language has. Thus, different combinations of symmetric and asymmetric negative constructions and paradigms are found in type SymAsy.

It is worth noting that not all asymmetry found in declarative verbal main clauses is taken into account in this classification. First, asymmetry that is not grammaticalized but is only a tendency does not figure on the maps; e.g. in Russian there is a preference for imperfective aspect under negation, but both perfective and imperfective aspect can appear under negation (see for example Schmid 1980). Russian is therefore classified as having only symmetric negation. Secondly, asymmetries that are outside the limits of the standard negation construction are disregarded. Since standard negation expresses the negation of verbal clauses, the negative construction is delimited in such a way that categories belonging to the verb or more globally to the clause are relevant, but categories belonging to individual nonverb constituents (like case marking on noun phrases) usually are not. For example, the changes in case in Finnish (accusative changing into partitive under negation) and the changes in the article in French (indefinite article changing into partitive *de* under negation) are not taken into account. They belong primarily to the marking of noun phrases and are not considered part of the negative construction here. Similarly, negative polarity effects on indefinites are not taken into account. Thirdly, changes that are (morpho)phonological and do not involve any substantial differences in the meaning-bearing elements are not relevant for the classification, either.

The map says nothing about the proportion of the different types in individual languages. In some cases the proportion can be rather unequal; standard negation in a language can be predominantly of a given type, but there can be a small domain where the language shows a different type. Nevertheless, the symbol given to the language represents both types equally strongly. Haitian Creole, for example, has a symmetric negative construction and most verbal categories can appear under negation (d'Ans 1968). But the particle *ava* marking future cannot occur in negated clauses. This asymmetry concerns only a minority of the negative clauses in the language, but still Haitian Creole is treated as belonging to Type SymAsy.

2 Geographical Distribution

Symmetric negation (Types Sym and SymAsy) is frequent in all parts of the world. There are two areas where only symmetric negation (Type Sym) is found: much of continental Europe and a large part of South-East Asia. Languages with only symmetric negation (Type Sym) are the least common in Africa, but on the whole symmetric negation is not uncommon in Africa, since a large number of African languages have both symmetric and asymmetric negation (Type SymAsy). Asymmetric negation (Types SymAsy and Asy) is also frequent in all parts of the world. It is absent in the two areas just mentioned and most common in Africa. Languages with only asymmetric negation (Type Asy) are not dominant in any area. Note that some connections between Maps 113 and 114 are discussed in Chapter 114.

3 Theoretical issues

Why have I made this division into symmetric and asymmetric negation? The main reason is that it allows us to find different functional motivations for the different types of negative structures. Symmetric negatives copy the formal linguistic structure of the affirmative, and are thus language-internally analogous to the affirmative structure; they are motivated by the pressure for cohesion in the system. As to asymmetric negation, there is functional (semantic and pragmatic, i.e. in a strict sense language-external) asymmetry between affirmation and negation: they differ as to their stativity/dynamicity, reality, and discourse context, for example (see Givón 1978, 2001: 369–98; Miestamo 2003). Asymmetric negative structures copy (or grammaticalize) aspects of this functional asymmetry, and are thus language-externally analogous to the just-mentioned functional asymmetry. The different subtypes of asymmetric negation discussed in Chapter 114 have grammaticalized different aspects of the functional asymmetry. For more detailed discussion, see Miestamo (2000, 2003: 169–212).

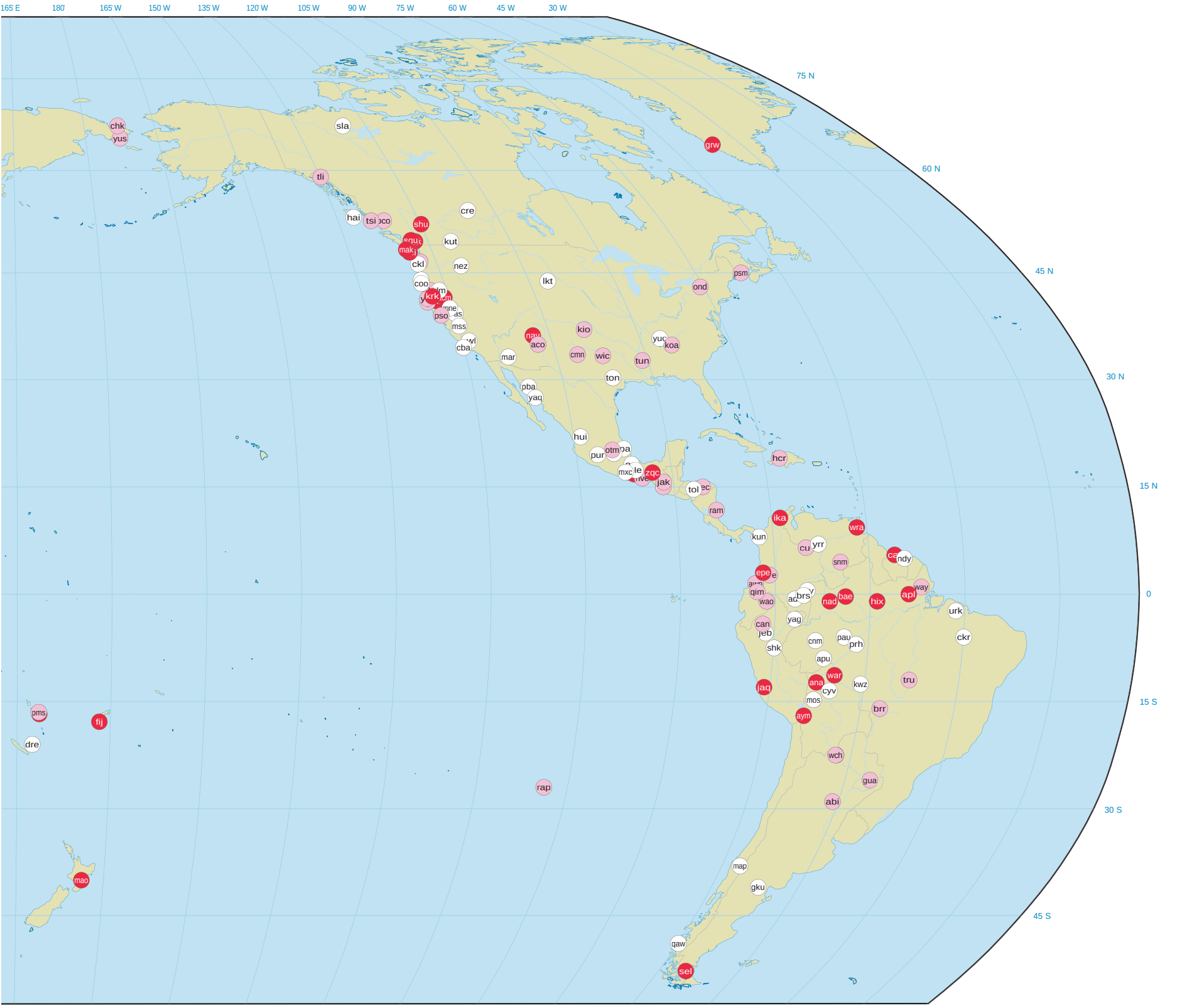
113 Symmetric and Asymmetric Standard Negation

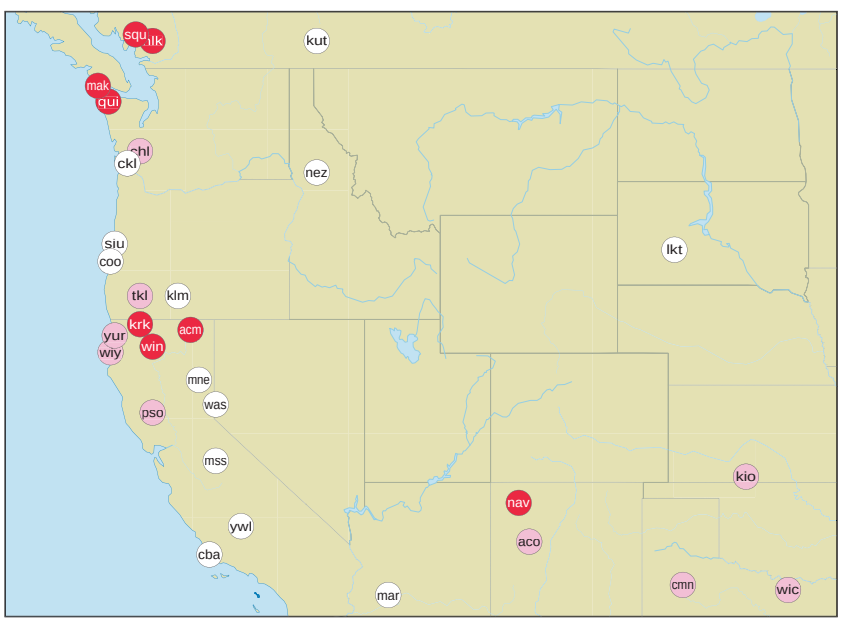
MATTI MIESTAMO

Symmetric [114]
Asymmetric [53]
Both [130]

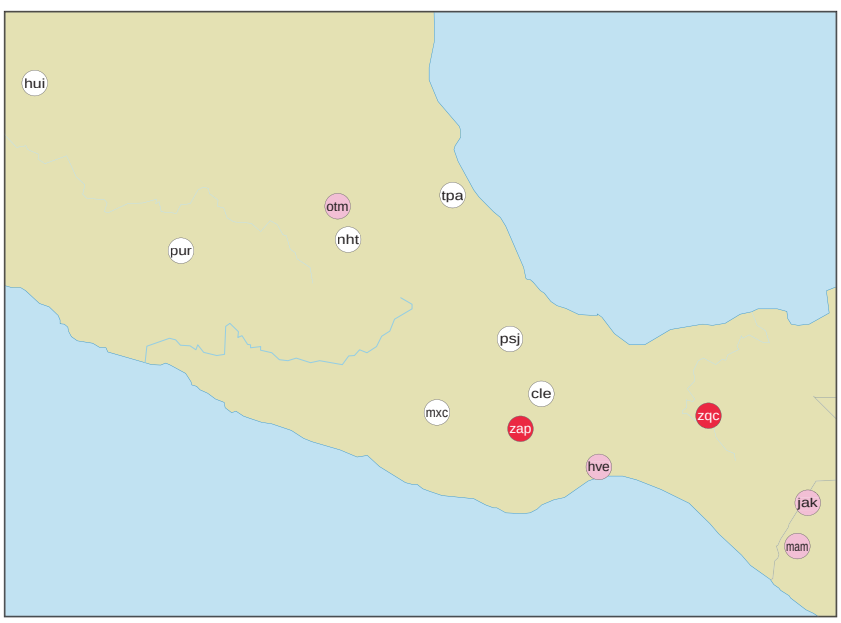
Nigeria and surrounding areas

New Guinea and northern Australia





western United States



southern Mexico

114 Subtypes of Asymmetric Standard Negation

MATTI MIESTAMO

1 Definition of values

The term **standard negation** refers to the basic way(s) a language has for negating declarative verbal main clauses. In **symmetric negation** affirmative and negative structures are identical except for the presence of the negative marker(s), whereas in **asymmetric negation** the structure of negatives differs from the structure of affirmatives in other ways too. Structural differences, i.e. asymmetry, can be found either between the affirmative and negative **constructions** or between the **paradigms** that the affirmative and negative constructions form. These distinctions are defined and illustrated in connection with Map 113, which shows the geographical distribution of symmetric and asymmetric standard negation.

Map 114 shows how the different subtypes of asymmetric standard negation are distributed across the languages of the world. The asymmetry between affirmation and negation can be found in different grammatical domains, and subtypes of asymmetric negation can be established according to the domain and nature of the asymmetry. Three subtypes are distinguished in the present context, according to whether the asymmetry is connected (i) to the **finiteness** of verbal elements (A/Fin), (ii) to the marking of **reality status** (A/NonReal), or (iii) to the marking of **verbal categories** in some other ways (A/Cat). As a language can have different negative structures in different contexts, different subtypes of asymmetric negation can also be found within one and the same language. The map therefore distinguishes three further types of language where two of the three subtypes of asymmetric negation are combined. Naturally, the subtype classification can only be applied to languages where asymmetric negation is found in the first place (Types Asy and SymAsy in Chapter 113); languages where only symmetric negation is found (Type Sym in Chapter 113) are thus irrelevant to the present map (number seven in the box).

The standard negation structure(s) found in a language can show asymmetry:

1.	In finiteness: Subtype A/Fin	40
2.	In reality status: Subtype A/NonReal	20
3.	In other grammatical categories: Subtype A/Cat	82
4.	In finiteness and reality status: Subtypes A/Fin and A/NonReal	9
5.	In finiteness and other grammatical categories: Subtypes A/Fin and A/Cat	21
6.	In reality status and other grammatical categories: Subtypes A/NonReal and A/Cat	11
7.	Non-assignable (no asymmetry found)	114
	total	297

Some of the three main subtypes can be divided into further subtypes (see Miestamo 2003). The theoretical significance of the classification is discussed briefly in Chapter 113.

In **Subtype A/Fin** negation affects the finiteness of verbal elements. Typically, the negative construction adds a new finite element (finite verb) to the clause, and the lexical verb becomes nonfinite and/or subordinate to the added finite element. A negative construction of this subtype is found in Hixkaryana (Carib; Brazil; see (1)).

(1) Hixkaryana (Derbyshire 1979: 48)

a. *ki-amryeki-no*
1.SUBJ-hunt-IMM.PST
'I went hunting.'

b. *amryeki-hira w-ah-ko*
hunt-NEG 1.SUBJ-be-IMM.PST
'I did not go hunting.'

In Hixkaryana a (non-negative) copula functions as the finite element of the negative clause, and the negative marker is a deverbalizing suffix on the lexical verb. Negative verb constructions, as found for example in Finnish (see Chapter 113, example (4)), also belong to Subtype A/Fin. A new finite element is not necessarily added in Subtype A/Fin, but the lexical verb still loses its finiteness (e.g. in Sentani; see Hartzler 1994).

Subtype A/NonReal is asymmetric with respect to the marking of reality status of events—the negative clause is obligatorily marked by a nonrealized category, whereas the affirmative is not. In Imbabura Quechua (see (2)), in addition to the negative particle *mana*, the negative clause contains the marker *-chu* that also appears in a nonrealized context, viz. yes/no questions.

(2) Imbabura Quechua (Cole 1985: 83, 94)

a. *juzi iskay kaballu-ta chari-n*
José two horse-ACC have-3
'José has two horses.'

b. *ñuka wawki mana jatun wasi-ta chari-n-chu*
my brother NEG big house-ACC have-3-NEG/Q
'My brother does not have a big house.'

c. *kan-paj wawki jatun wasi-ta chari-n-chu*
you-POSS brother big house-ACC have-3-NEG/Q
'Does your brother have a big house?'

The negative declarative is symmetric with the non-negative interrogative, since they only differ by the presence of the negative *mana*. The A/NonReal asymmetry is paradigmatic. The distinction between interrogative and declarative is not lost in negatives since there is a separate construction for negative interrogatives. Most commonly the nonrealized category involved in A/NonReal asymmetry is a more general irrealis category; e.g. in Maung (see Chapter 113, example (7)), where the realis-irrealis distinction made in the affirmative is lost in the negative, resulting in paradigmatic A/NonReal asymmetry.

Subtype A/Cat involves changes in the marking of grammatical categories (such as tense, aspect, mood, person, number, etc.) under negation. For example, in Karok (Hokan; California; see (3)) the affirmative and the negative use different person-number affixes, and in Koyraboro Senni (Songhay; Mali; see (4)) the negative imperfective marker replaces the positive imperfective marker.

(3) Karok (Bright 1957: 67)

a. *kun-iykár-at*
3PL>3SG-kill-PST
'They killed [him/her].'

b. *pu-ʔiykar-áp-at*
NEG-kill-3PL>3SG-PST
'They did not kill [him/her].'

(4) Koyraboro Senni Songhay (Heath 1999a: 8–9)

a. *n ga koy*
2SG.SUBJ IMPF go
'You (SG) are going.'

b. *war si koy*
2PL.SUBJ NEG.IMPF go
'You (PL) are not going.'

The constructional asymmetry in Lezgian past imperfectives (see Chapter 113, example (6)) is also a case of A/Cat asymmetry. In Páez (Paezan; Colombia; see (5)) the negative construction is symmetric but the paradigm is asymmetric.

(5) Páez (Jung 1989: 102–4)

a. *âts,h-a' ts,hab-na u'x-we-ts-thu*
now-TOP village-to go-IMPF-PROG-DECL.1SG
'I'm going to the village right now.'

b. *u'x-we-ts-me:-th*
go-IMPF-PROG-NEG-DECL.1SG
'I don't go.' / ['I'm not going.']

c. *skwela-na-t, u'x-we-'*
school-to-FACT.3PL go-IMPF-HAB
'They go to school.'

d. **u'x-we-'-me:-th*
go-IMPF-HAB-NEG-DECL.1SG

As can be seen in (5b), the negative suffix *-me:* is simply added to the corresponding positive (5a) with no further structural change (the loss of the final vowel is morphophonological). But the habitual cannot appear in the negative (see 5c, d) and the distinction between the progressive and the habitual is neutralized. This is a case of paradigmatic A/Cat asymmetry. In Burmese (see Chapter 113, example (5)) the affirmative paradigm makes a distinction between actual, potential (future) and perfect, but as the negative construction replaces the tense-aspect-mood marker with the negative suffix *-bû,* the distinction is lost, and both the construction and the paradigm show A/Cat asymmetry. In Swahili the paradigm is thoroughly asymmetric, since affirmatives and negatives use different sets of tense-aspect markers (see Contini-Morava 1989).

Sometimes a construction involving an auxiliary or an auxiliary-like element in the positive is negated by the replacement of the auxiliary by a corresponding irregular/suppletive negative auxiliary. The asymmetry is derived from the asymmetry in the negation of the auxiliary, and the idiosyncrasy in the negation of this particular element is not usually relevant to standard negation. But in cases where the auxiliary is used in a productive periphrastic verbal construction, the asymmetry is relevant. These constructions belong to Subtype A/Cat. In Ladakhi (Tibeto-Burman; Kashmir; see (6)) the auxiliaries *yin* and *yod* jointly mark present continuous tense, but in the negative the latter is replaced by its negative form *med*.

(6) Ladakhi (Koshal 1979: 238)

a. *pəlldənni* *ʂpečhə* *ɖi-yin-yot*
Paldan.ERG book.ABS write-AUX-AUX
'Paldan is writing a book.'

b. *pəlldənni* *ʂpečhə* *ɖi-yin-met*
Paldan.ERG book.ABS write-AUX-NEG.AUX
'Paldan is not writing a book.'

A similar case is found in Latvian third-person periphrastic tenses.

As we have seen, the marking of grammatical categories changes under negation in all three of the subtypes. But in Subtype A/Cat no generalizations can be made about the functional effects of the changes as can be done in the two previous types—changes in finiteness in A/Fin and nonrealized marking in A/NonReal. The negative past imperfective marker in Lezgian and the negative person-number markers in Karok neither change the finiteness of the verb, nor mark irrealis or any other nonrealized categories. They are merely variants appearing under negation with no further generalizations made about their functions elsewhere in the grammars of these languages. Sometimes A/Cat asymmetry does have an identifiable semantic effect, e.g. when distinctions made in the affirmative are neutralized in the negative, but these effects do not lend themselves to relevant cross-linguistic generalizations and give no grounds for distinguishing new subtypes of asymmetric negation; note that perfectives or completives do not show a tendency to be excluded in negatives, as is often claimed (e.g. Schmid 1980, see Miestamo 2003: 149–50 for discussion). There is, however, one case where generalizations can be made: there are languages where negatives use a form that expresses emphasis in non-negatives; these structures are treated as an independent (although marginal) subtype of asymmetric negation (A/Emph) in Miestamo (2003). The negation of simple tenses in English has A/Emph asymmetry, as the distinction between emphatic and non-emphatic is lost in the negative (*Chris reads* versus *Chris does read* versus *Chris does not read*). A/Emph asymmetry is found in only a few languages and is treated here as A/Cat, since too many distinctions cannot be made on the maps. Some languages show asymmetries that cannot easily be connected to the marking of any category, e.g. the order of auxiliary and main verb changes in negatives in contemporary standard Basque (Saltarelli et al. 1988; José Hualde, p.c.); such asymmetries are found in only a few languages and are treated as A/Cat here.

Due to space limitations, it is not possible to give examples from languages with multiple subtypes. I will, however, mention one or two languages belonging to each category.

Subtypes A/Fin and A/NonReal: In Copainalá Zoque (Mixe-Zoque; Chiapas, Mexico; Harrison et al. 1981) the non-past uses a negative verb construction (A/Fin), whereas the past negative construction contains a suffix that also marks some irrealis categories (A/NonReal). In Squamish (Salishan; British Columbia; Kuipers 1967) the negative construction has both A/Fin asymmetry in that the negative element is a verb, the negated content being expressed in a subordinate clause, and A/NonReal asymmetry in that the subordinate clause is introduced by the irrealis modal clitic.

Subtypes A/Fin and A/Cat: In Kolokuma Ijo (Ijoid, Niger-Congo; Nigeria; Williamson 1965) a negative copula appears as the finite element in future negation, which is thus of Subtype A/Fin; many other tense-aspects show A/Cat asymmetry: there are changes in how the tense-aspect categories are marked, and the paradigm also shows A/Cat asymmetry.

Subtypes A/NonReal and A/Cat: In Warndarang (Maran; Northern Territory, Australia; Heath 1980) past and present negatives are irrealis-marked (versus realis-marked affirmatives) (A/NonReal), and in the future the negator replaces the potential (future) prefix (A/Cat).

Note that as was the case in Chapter 113, the map has nothing to say about the prominence of the different types in individual languages. The combination of all three subtypes was found in two languages, Imonda (Border family; Papua New Guinea) and Nivkh (isolate; Russian Far East), but for the sake of simplicity, no separate symbol was given to this combination—the map shows Imonda and Nivkh as having A/Fin and A/NonReal asymmetry.

Sometimes a language seems to have asymmetry of two different subtypes, but on closer examination, one of these asymmetries turns out to be derived from the other. In these cases the derived asymmetry is not represented on the map. In Sentani (Sentani family; Papua, Indonesia) the lexical verb loses all verbal marking and thus loses its finiteness. From this A/Fin asymmetry it follows that all distinctions made in the verb are lost, and the paradigm seems to have A/Cat asymmetry. But the paradigmatic asymmetry is subordinate to the A/Fin asymmetry and hence is not taken into account. In Maung (see Chapter 113, example (7)) the punctual-continuous distinction is lost in the negative, suggesting A/Cat asymmetry in the paradigm. It is, however, the irrealis category which is responsible for the loss of this distinction. Irrealis verbs cannot be distinguished for punctual/continuous, and since the negative is irrealis-marked, the distinction is lost in all negatives too.

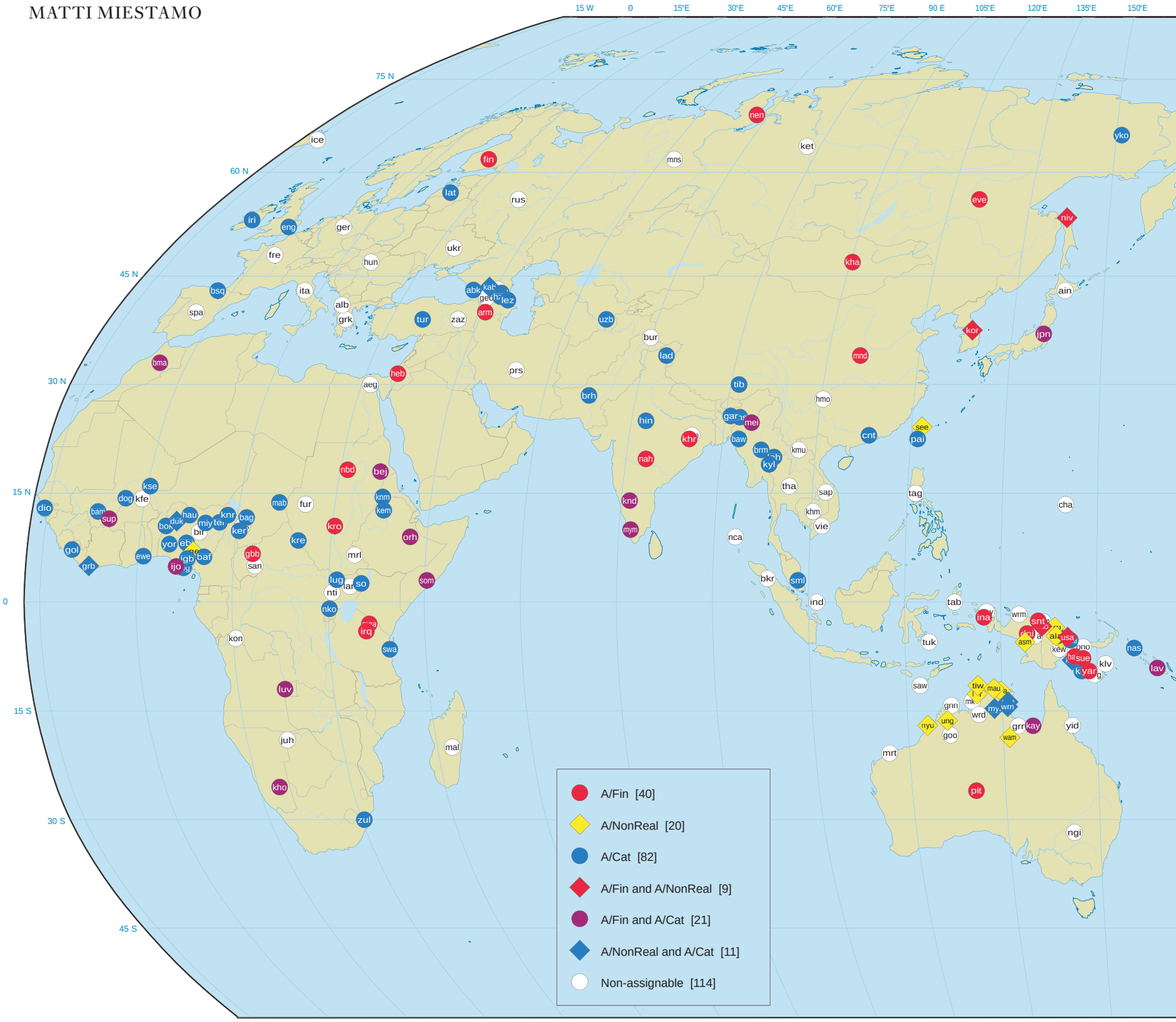
2 Geographical Distribution

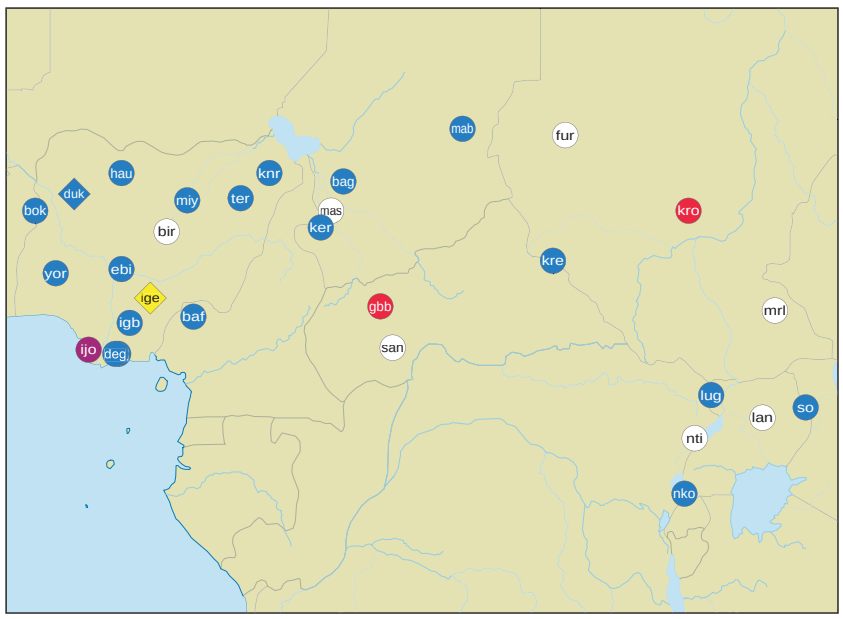
Subtype A/Fin asymmetry is found in most parts of the world; the strongest areas are in northern and eastern Eurasia, South America, New Guinea, and northeastern Africa; a local hotbed is found in the northwestern United States and southwestern Canada. Subtype A/NonReal is frequent in Australia, and some cases are also found elsewhere in the Pacific Rim; in other parts of the world this subtype is rare. Subtype A/Cat is common almost everywhere in the world; it is very frequent in Africa and frequent in North America and South Asia as well; the Caucasus presents a local hotbed of this subtype.

Some connections between Maps 113 and 114 deserve to be pointed out. In the areas where A/Cat asymmetry is common, Type SymAsy is also very common; this is due to the circumstance that often A/Cat asymmetry only affects some verbal categories in a language and negation is symmetric elsewhere, or that asymmetry is only found in the paradigm while the negative construction is symmetric. Many of the languages with A/NonReal asymmetry are also languages of Type SymAsy; in A/NonReal asymmetry the construction is often symmetric, the asymmetry being in the paradigm. Languages that have only Subtype A/Fin asymmetry often belong to type Asy; this is explained by the fact that A/Fin asymmetry is almost always constructional, not paradigmatic (paradigmatic asymmetry is in principle irrelevant to type Asy).

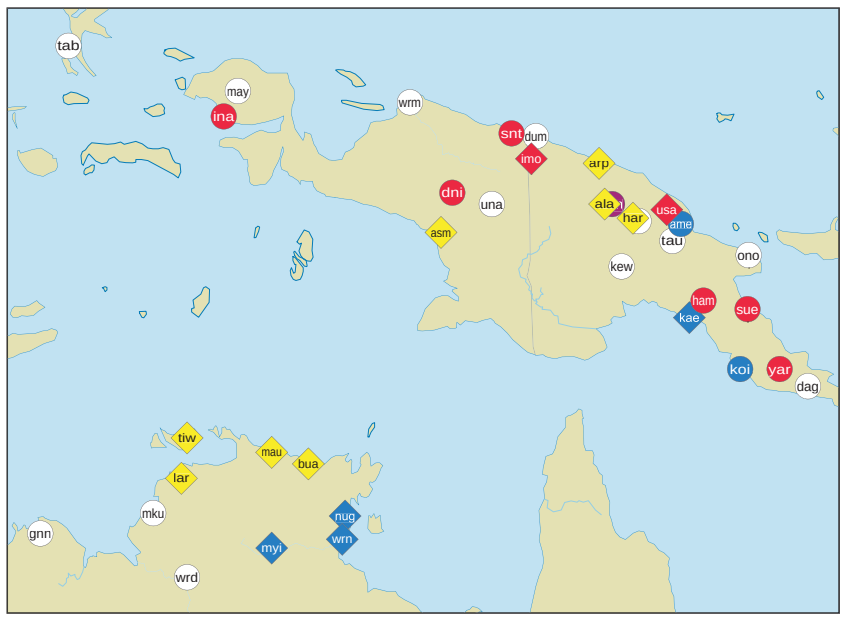
114 Subtypes of Asymmetric Standard Negation

MATTI MIESTAMO

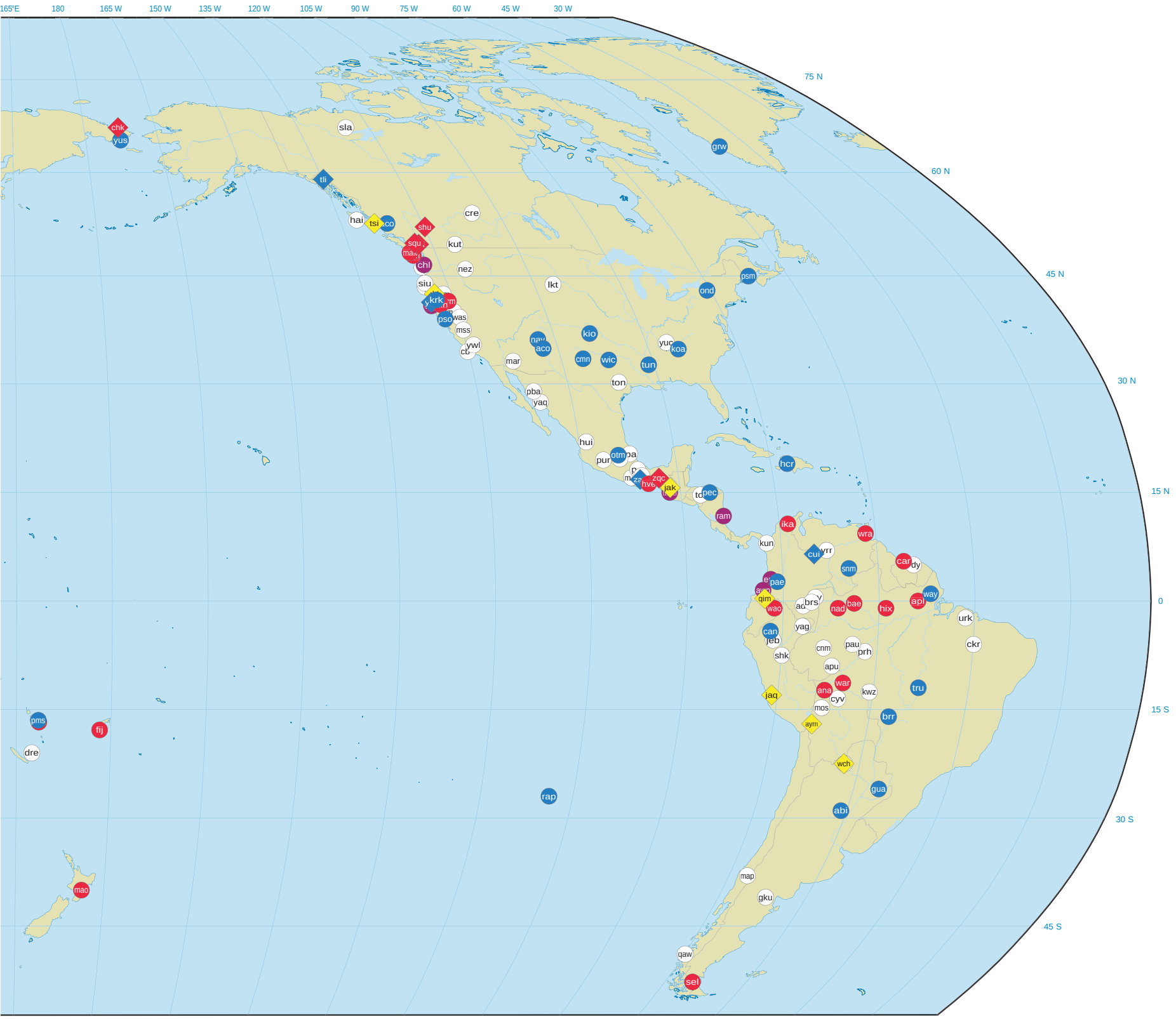


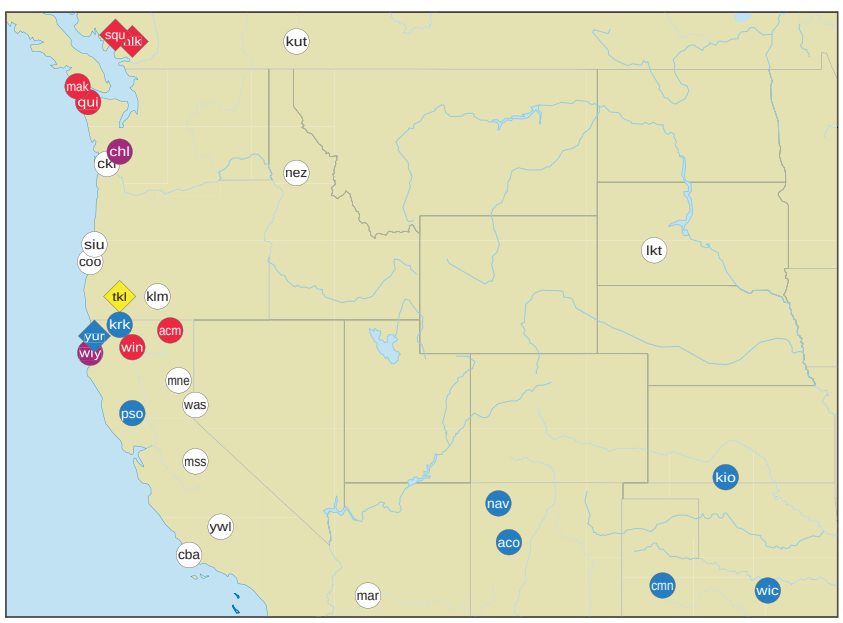


central Africa

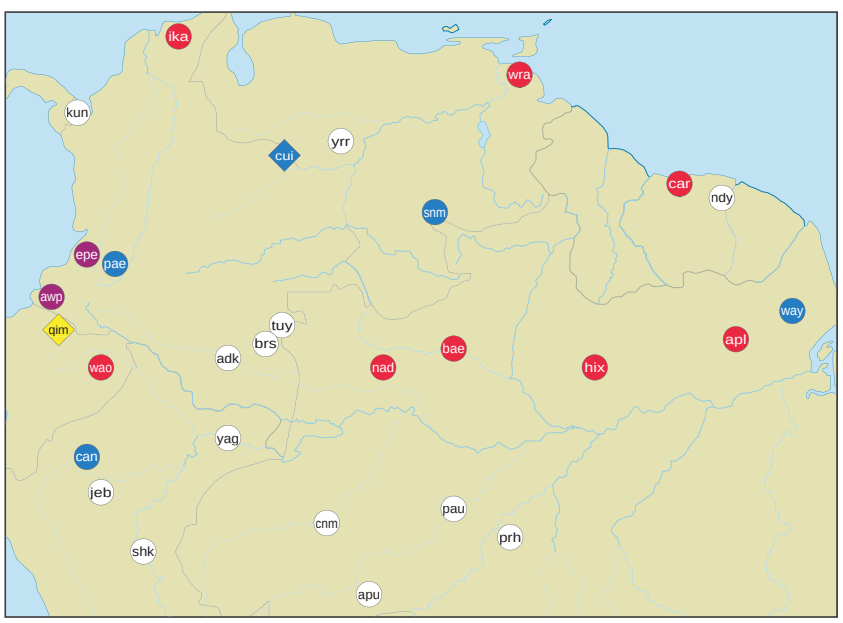


New Guinea and northern Australia





western United States



northern South America

115 Negative Indefinite Pronouns and Predicate Negation

MARTIN HASPELMATH

1 Overview of the values

Negative indefinite pronouns like 'nobody', 'nothing', 'nowhere' may or may not co-occur with the ordinary marker of predicate negation. A language where they **always co-occur with predicate negation** is Russian. In (1a–b), omission of the predicate negator *ne* would render the sentences ungrammatical.

(1) Russian

a. ***Nikto*** *ne* *prišel.*
nobody NEG came.
'Nobody came.'

b. *Ja* *ne* *videla* ***ničego.***
I NEG saw nothing
'I saw nothing.'

A language where negative indefinites **never co-occur with predicate negation** is German. In (2a–b), use of the predicate negator *nicht* would render the sentences ungrammatical (**Niemand kam nicht; *Ich habe nichts nicht gesehen*).

(2) German

a. ***Niemand*** *kam.*
nobody came

b. *Ich* *habe* ***nichts*** *gesehen.*
I have nothing seen

Languages may also show **mixed behaviour**. For example, Spanish negative indefinites preclude predicate negation when they precede the verb (as in 3a), but predicate negation is required when they follow the verb (as in 3b).

(3) Spanish

a. ***Nadie*** *vino.*
nobody came

b. *No* *vi* ***nada.***
NEG I.saw nothing

Finally, some languages require a **negative existential construction** with a negative or negated existential verb as the main predicate, and the word translating the English indefinite pronoun as the subject. An example is Nêlêmwa (Oceanic; New Caledonia):

(4) Nêlêmwa (Bril 1999: 84)
Kia *agu* *i* *uya.*
not.exist person 3SG arrive
'Nobody came.' (Lit. 'There isn't a person who came.')

One could say that these languages, too, show co-occurrence with predicate negation, but they are assigned to a separate category here because these languages do not allow their counterparts of negative indefinites to occur in the standard nominal or adverbial positions.

Thus, four values are distinguished on the map:

1.	Negative indefinites co-occur with predicate negation	170
2.	Negative indefinites preclude predicate negation	11
3.	Negative indefinites show mixed behaviour	13
4.	Negative existential construction	12
	total	206

2 What counts as a negative indefinite pronoun

Negative indefinite pronouns are defined as nominal or adverbial expressions that directly translate 'nobody', 'nothing', 'nowhere', 'never' (etc.) in sentences (1)–(3), independently of whether they co-occur with predicate negation. They fall into the same formal types as the indefinite pronouns that are discussed in Chapter 46:

(i) interrogative-based negative indefinites;
(ii) generic-noun-based negative indefinites; and
(iii) special negative indefinites.

Typically, languages with interrogative-based specific indefinites also have interrogative-based negative indefinites, e.g. Georgian *vin* 'who', *vi-ɣac* 'somebody', *ara-vin* 'nobody' (Hewitt 1995: 81–2), and languages with generic-noun-based specific indefinites also have generic-noun-based negative indefinites, e.g. Kana (Benue-Congo; Nigeria) *nū* 'thing; something; nothing' (Ikoro 1996: 345).

Note that the definition of **negative indefinite pronoun** that is used here does not make reference to any kind of "inherent" negative sense. Some linguists feel that English *nobody* is "inherently negative", while *anybody*, which co-occurs with predicate negation (as in *I did not see anybody*), is not "inherently negative" (e.g. Bernini and Ramat 1996: 118–20). Such a claim would seem to be even more plausible with items such as Kana *nū*, which probably simply means 'thing' and does not even have a special sense in (5), where it occurs in the scope of negation.

(5) Kana (Ikoro 1996: 345)
m̀-síì *yāɛ̄* ***nū.***
I-NEG buy thing
'I have not bought anything.'

However, in this chapter no such distinction between two types "inherently negative" and "not inherently negative" indefinites is made, because it is unclear which criteria should be used to identify them systematically. One criterion, the occurrence in elliptical negative answers (*Who did you see? Nobody/*Anybody*), is not practical for a large-scale cross-linguistic study because of insufficient data, and it does not always coincide with another possible criterion, the non-occurrence in non-negative contexts. For instance, Modern Greek *kanénas* 'anybody, nobody' can be used in elliptical negative answers (unlike English *anybody*), but it is widely used in nonnegative contexts (like *anybody*).

Thus, for this chapter all nominal expressions that correspond to 'nobody' and 'nothing' are regarded as "negative indefinite pronouns", even though in many or most languages the negative sense is contributed exclusively by the predicate negation (see Haspelmath 1997: 193–200 for further discussion).

3 What counts as predicate negation

For a language to qualify as belonging to the first type (co-occurrence with predicate negation), the negator must be the same as in ordinary negative clauses lacking a negative indefinite pronoun. This is indeed by far the most common situation, illustrated in (6) from Abun (West Papuan; Papua, Indonesia). This language has bipartitite predicate negation (*yo . . . nde*).

(6) Abun (Berry and Berry 1999: 131, 143)

a. *Isak* *yo* *ben* ***suk*** *mo* *nggwe* *nde.*
Isak NEG do thing LOC garden NEG
'Isak did not do anything at the garden.'

b. *Án* *yo* *ma* *mo* *nu* *nde.*
they NEG come LOC house NEG
'They did not come to the house.'

In most cases, there is no danger of mistaking the predicate negator for a part of the indefinite, because the negator has a fixed position in the clause, generally close to the verb (see Dryer 1988b). However, sometimes the position of the negator is independent of the position of the verb, as in Gooniyandi (Bunaban; Western Australia):

(7) Gooniyandi (McGregor 1990: 479, 389)

a. *Mangaddi* ***ngoorndoo-ga-ngaddaya*** *wardnginbidda.*
NEG who-ERG-INDEF they.took.me
'No one gave me a lift.'

b. *Mangaddi* *yangbala-ngga* *goornboo* *doownga*.
NEG young.man-ERG woman he.got.her
'Young men didn't take wives.'

In (7a), one might conceivably interpret *mangaddi ngoorndoo-ga-ngaddaya* as a single negative indefinite that does not co-occur with predicate negation; but as (7b) illustrates, Gooniyandi's negator typically occurs initially and need not be adjacent to the verb. Thus, *mangaddi* in (7a) is just the ordinary predicate negation, and we have the same negative construction in (7a) and (7b).

The situation is different in Tzutujil (Mayan; Guatemala):

(8) Tzutujil (Dayley 1985: 245, 321)

a. ***Majuun*** *x-in-tz'at*.
nothing COMP-1SG.ERG-see
'I saw nothing.'

b. *Ma* *x-b'e* *ta* *jar* *Aa* *Lu7*.
NEG COMP-go IRREAL the young Pedro
'Pedro did not go.'

In (8a) *ma-* in *ma-juun* [NEG-one] does not count as predicate negation, because (i) the predicate negator *ma* is generally immediately preverbal, as in (8b); (ii) it is a separate word rather than a prefix, and (iii) it is generally accompanied by the irrealis particle *ta* in the completive aspect. Thus, (8a) represents a different negative construction from (8b).

Besides clear cases like Gooniyandi and Tzutujil, there are also truly doubtful cases like Chalcatongo Mixtec (Oto-Manguean; Mexico). In this language, predicate negation is ordinarily expressed by the pre-predicate clitic *tu*=, as in (9a). However, this clitic may also precede a focused noun phrase, as in (9b), so that it is possible that we are dealing with the same construction in (9c). If this were the case, Chalcatongo Mixtec would be classified as type one.

(9) Chalcatongo Mixtec (Macaulay 1996: 120, 124)

a. *Ndežu* *tú=žaʔu*.
food NEG=expensive
'The food is not expensive.'

b. *Tú=šūʔū* *ñabaʔa=rí*.
NEG=money have=1
'I didn't have any money.'

c. ***Tú=ndéú*** *ní=kii*.
NEG=who COMP=come
'Nobody came.'

However, I have classified it as type two (precluding predicate negation), because *ndéú* 'who' does not seem to be otherwise used in an indefinite sense, and in the corresponding *tú=kwɨtɨ* 'nothing', the element *kʷɨtɨ* is unique. Thus, *tú=ndéú* and *tú=kʷɨtɨ* seem to be non-compositional expressions, but ideally one would have to have more information to be able to confirm this analysis.

4 Types of mixed behaviour

There are several different ways in which a language can show mixed behaviour. One type is the Spanish type, where the use of the predicate negator depends on the position of the indefinite with respect to the verb (see (3) above). Another type is represented by Swedish, which has two different sets of indefinites, one of which co-occurs with predicate negation, while the other precludes it (as the translations show, English is mixed in the same way):

(10) Swedish

a. *Jag* *vet* *inte* ***något***.
I know NEG something
'I do not know anything.'

b. *Jag* *vet* ***ingenting***.
I know nothing
'I know nothing.'

Third, a language may have a bipartite negator, only one of whose parts is precluded by the negative indefinite. This is the case in Maltese, whose ordinary predicate negator consists of the preverbal particle *ma* and the postverbal suffix *-x*. When the clause contains a negative indefinite, only *ma* appears (French, with its bipartitite negator *ne . . . pas*, behaves in the same way).

(11) Maltese (Haspelmath 1997: 200)

a. *Patricia* *ma* *rat-x* *lit-tifel*.
Patricia NEG she.saw-NEG ACC.the-boy
'Patricia did not see the boy.'

b. *Patricia* *ma* *rat(*-x)* ***xejn***.
Patricia NEG saw nothing
'Patricia did not see anything.'

5 The issue of "double negation"

In a number of Germanic languages whose normative varieties have indefinites precluding predicate negation, these indefinites do co-occur with negation in nonstandard varieties. For example, in Bavarian German, structures like (12) are completely normal (see, for example, Weiß 1999).

(12) Bavarian German
Gesdan *han* *e* ***neamd*** *ned* *gseng*.
yesterday have I nobody not seen
'Yesterday I saw nobody.'

Such sentences are often criticized from a normative point of view as being "illogical", because according to logic, two negations should cancel each other out. Two problems with this point of view are (i) that the negative indefinites in question cannot easily be characterized as "inherently negative" because this concept is difficult to define (see §2 above), and (ii) that the co-occurrence of negative indefinites with predicate negation, as in (12), is by far the most common type in the world's languages. One would not, after all, want to say that only a few Germanic languages are "logical". In fact, it seems that ultimately the normative criticism of "double negation" is not based on logic, but on the prestige of Latin, which served as the model for the standard varieties of several western European languages, and which had negative indefinites precluding predicate negation (cf. Haspelmath 1997: §8.2).

Although it is not possible to say whether a negative indefinite is "inherently negative", we can ask whether it contains a morpheme that functions as a negator elsewhere in the language. And here an interesting generalization emerges: in the great majority of languages where negative indefinites preclude predicate negation, the indefinites themselves contain a negation marker. Some examples are given in (13).

(13)	Ossetic	*ma-cy*	[NEG-what]	'nothing'
	Chamorro	*ni hayi*	[NEG who]	'nobody'
	Tzutujil	*ma-juun*	[NEG-one]	'nobody'
	Chalcatongo Mixtec	*tú=ndéú*	[NEG=who]	'nobody'
	Dutch	*n-iets*	[NEG-something]	'nothing'

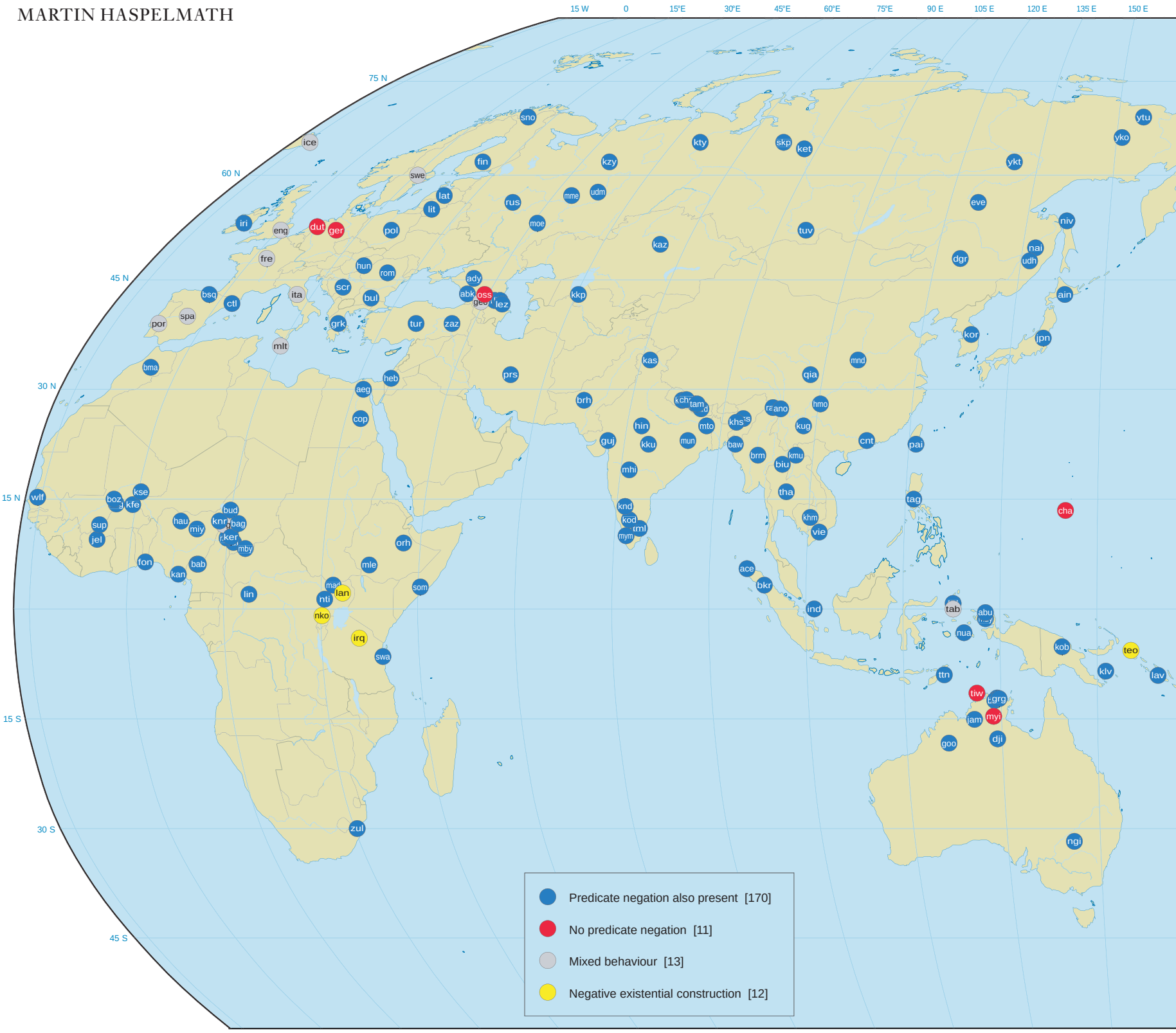
Exceptions to this generalization are found exclusively in western European languages (e.g. German *kein* 'no', Swedish *ingen* 'no(body)', and several Romance expressions which preclude verbal negation only when occurring preverbally, such as Spanish *nada* 'nothing', Italian *mai* 'never'). The probable diachronic explanation for this is that the pattern where negative indefinites preclude predicate negation typically arises by "attraction" of the negator to the negative indefinite pronoun (see (9) above for a possible intermediate stage in this process).

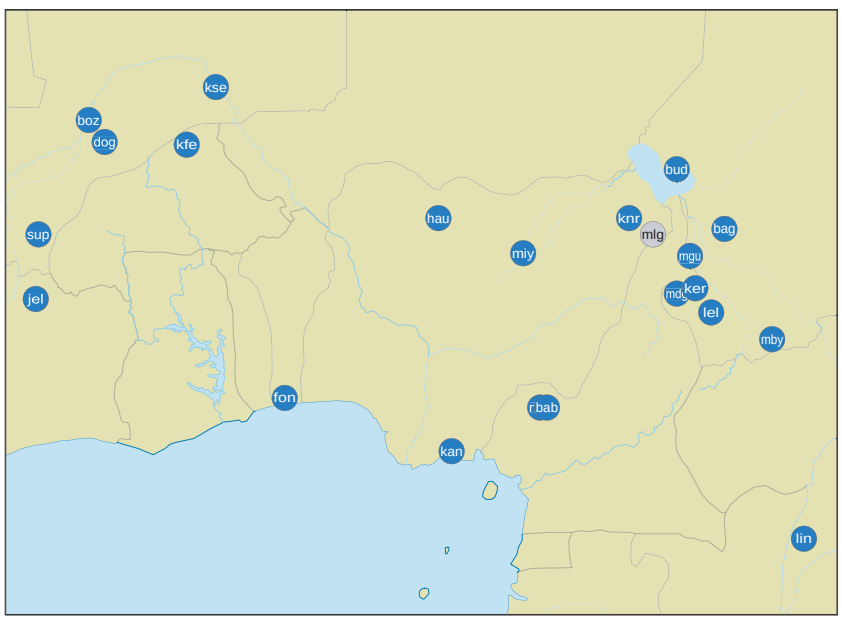
6 Geographical distribution

Negative indefinites precluding predicate negation are concentrated in Western Europe and in Mesoamerica, and languages with existential constructions are found particularly among the Oceanic languages of the small Pacific islands. The great majority of the world's languages have negative indefinites co-occurring with predicate negation.

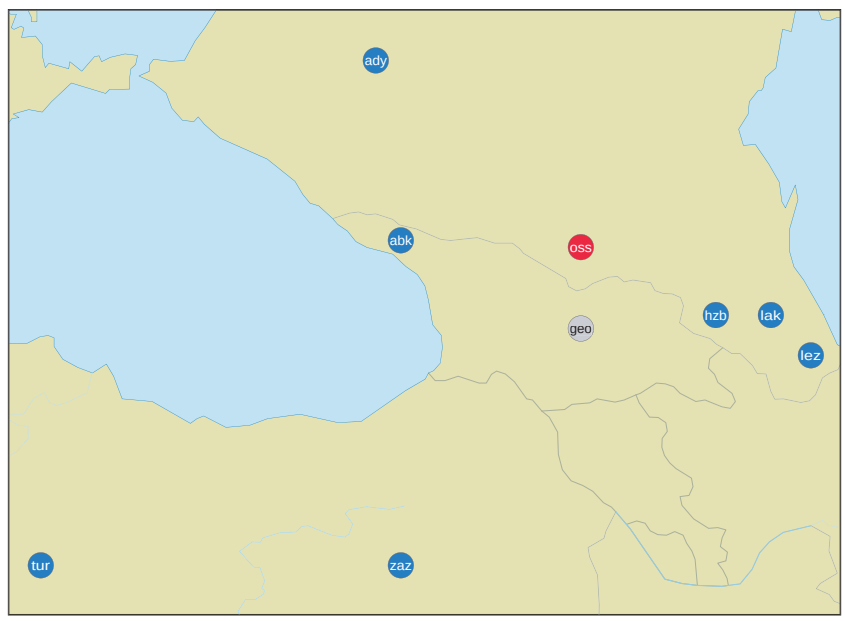
115 Negative Indefinite Pronouns and Predicate Negation

MARTIN HASPELMATH

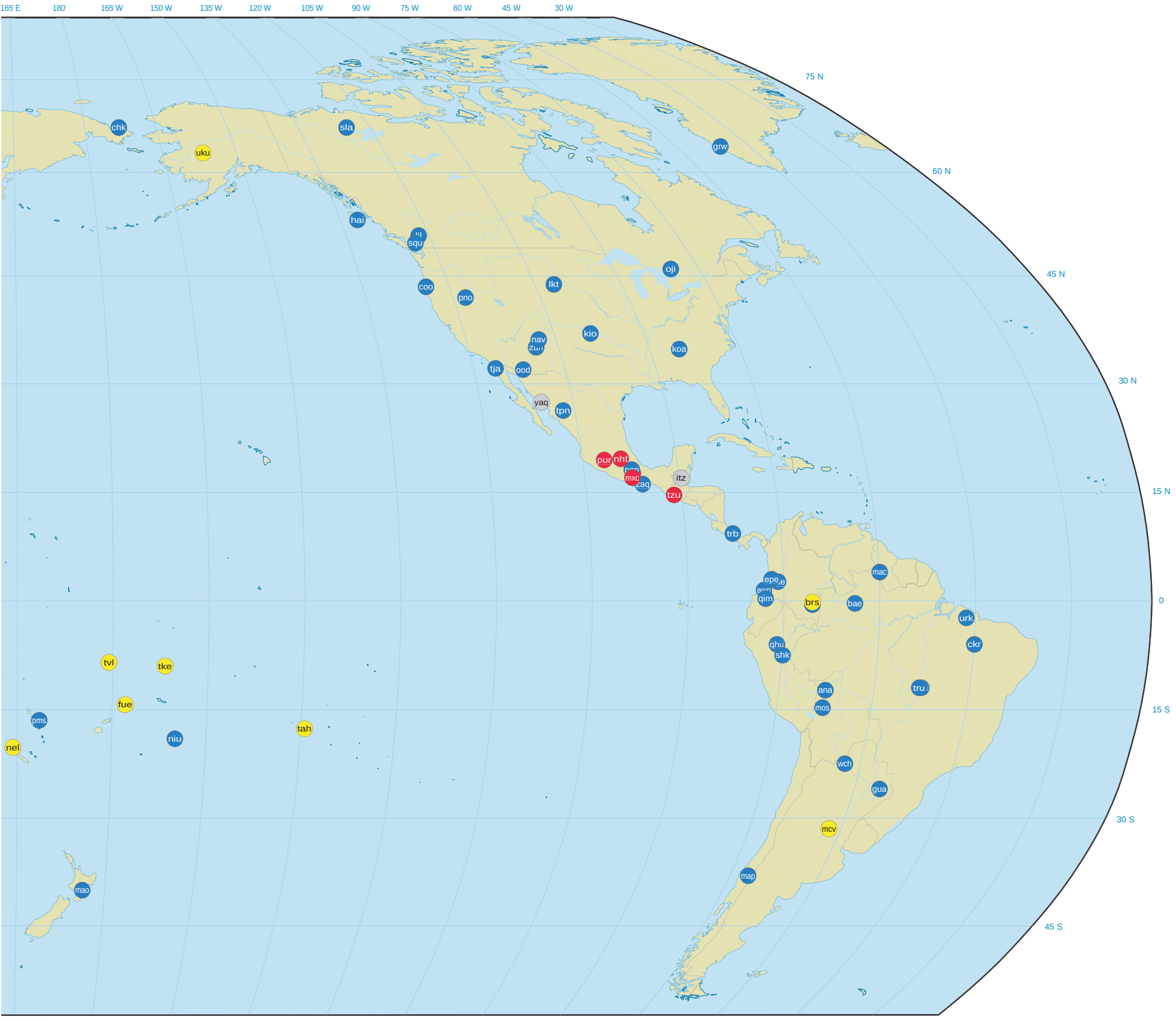


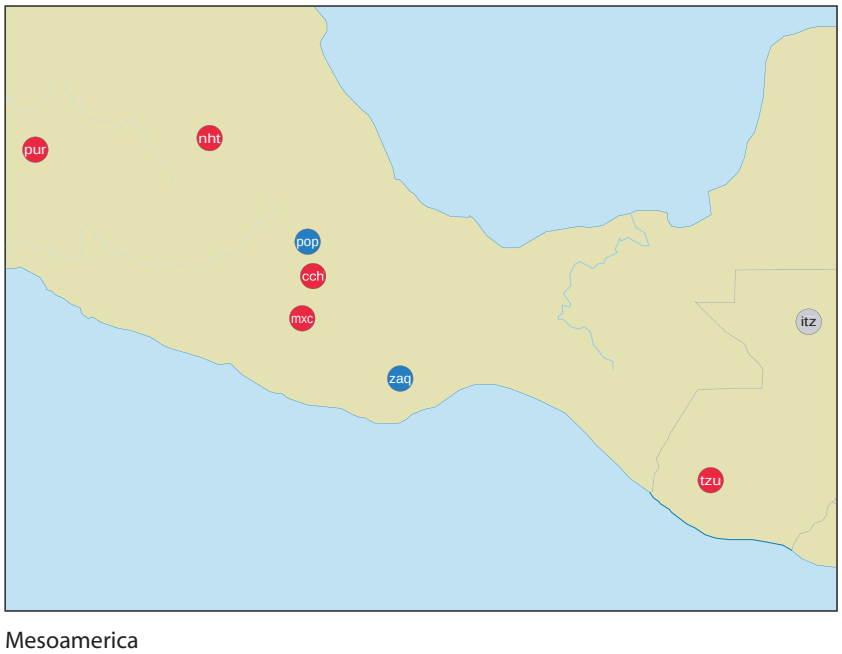


Nigeria and surrounding areas

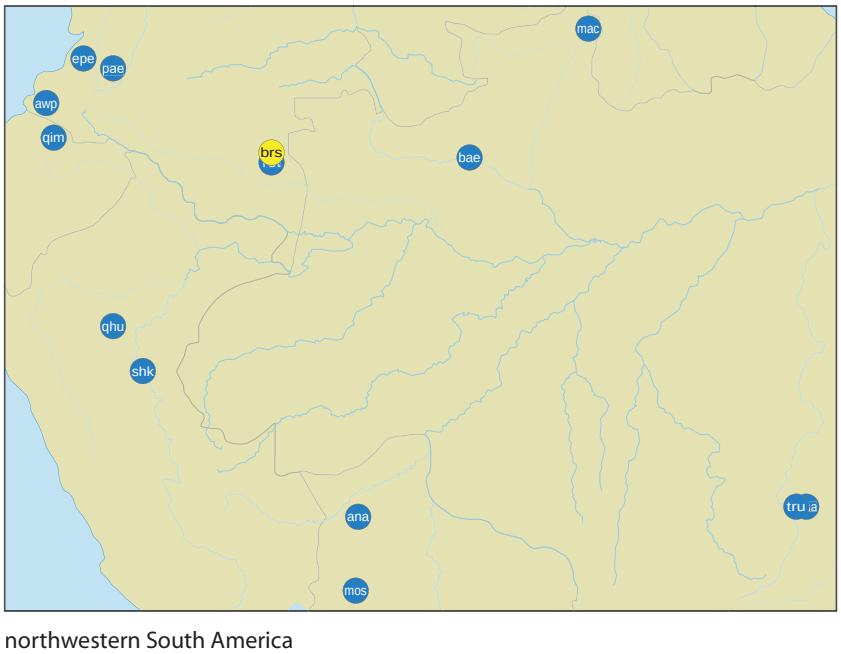


Caucasus and Black Sea area





Mesoamerica



northwestern South America

116 Polar Questions

MATTHEW S. DRYER

1 Defining the values

This map shows the method a language uses to indicate that an utterance is a **polar question**. Polar questions are ones to which the expected answer is the equivalent of "yes" or "no" (and which are thus sometimes called "yes-no questions"). They contrast with **content questions**, which contain an interrogative word meaning something like "who" or "where", in which some more specific answer is expected; content questions in English are often called "wh-questions". See Map 93 on content questions.

1.	Question particle	520
2.	Interrogative verb morphology	155
3.	Question particle and interrogative verb morphology	12
4.	Interrogative word order	12
5.	Absence of declarative morphemes	4
6.	Interrogative intonation only	138
7.	No interrogative–declarative distinction	1
	total	842

The first strategy for forming polar questions is the use of a **question particle** which is added to a corresponding declarative sentence to indicate that it is a question. In Maybrat (West Papuan; Papua, Indonesia), the interrogative morpheme is a particle that is added to the end of the sentence, as in (1).

(1) Maybrat (Dol 1999: 200)
ana *m-amo* *Kumurkek* ***a***
3PL 3-go Kumurkek Q
'Are they going to Kumurkek?'

In Kiowa (Kiowa-Tanoan; Kansas and Oklahoma), the question particle occurs at the beginning of the sentence, as in (2).

(2) Kiowa (Watkins 1984: 211)
hɔ́ *á-k'í·* *àn* *dét-mɔ́nyáygɔ́p* . . .
Q 2.POSS-husband HAB 2SG.OBJ-wave.IMPF
'Does your husband wave to you . . . ?'

Map 92 shows the distribution of different positions in which question particles occur. Because Chapter 92 provides more detailed discussion of question particles, discussion in this chapter is rather brief and the reader is referred to Chapter 92 for further discussion of question particles.

Interrogative **clitics**, which attach to some word, but which exhibit freedom as to the category of word they attach to, are treated here as question particles. For example, in Fyem (Platoid, Niger-Congo; Nigeria), the question particle attaches phonologically as a clitic to whatever is otherwise the last word in the sentence, as illustrated in (3), in which the question particle *=a* attaches to the object noun phrase *aré=n* 'the clothes'.

(3) Fyem (Nettle 1998: 50)
taa *wɔn* *aré=n=**a***
3SG.PERF wash clothes=DEF=Q
'Did she wash the clothes?'

The second general strategy for signalling polar questions involves the use of distinct **interrogative verbal morphology**. Most commonly, the verbal morphology may involve an affix that specifically signals that the utterance is a question, as in (4a) from Hunzib (Daghestanian; Russia) and (4b) from Tunica (isolate; Mississippi).

(4) a. Hunzib (van den Berg 1995: 112)
*ẽχ'e-čó-**y***
go-PRES.1/2-Q
'Are you going?'

b. Tunica (Haas 1940: 118)
lɔ'ta *wi-wa'nă-**n***
run 2SG-want-Q
'Do you want to run?'

In some languages, the verbal morphology involved in questions is more fusional and it is more difficult to isolate a specific morpheme as the interrogative morpheme. For example, in Gimira (Omotic, Afro-Asiatic; Ethiopia), there is a set of pronominal subject suffixes used in polar questions that are completely different from the subject affixes used in declarative sentences (Breeze 1990: 33). The distinction between separate question particles and interrogative affixes is not always clear. It is common in verb-final languages for question particles to immediately follow the verb, and they often loosely attach to the verb as clitics. Because it is difficult to distinguish verbal suffixes from clitics that can only attach to the verb, these are not distinguished here and morphemes which attach to the verb are all treated as instances of interrogative verb morphology, even if by some criteria they might be analysed as clitics.

The third type is languages that have **both question particles and interrogative verb morphology**, either as two separate constructions or occurring together in a single construction. For example, in Pirahã (Mura; Brazil), there is both a question particle, as in (5a), and an interrogative verbal suffix, as in (5b); it is also possible to use them together.

(5) Pirahã (Everett 1986: 236, 237)

a. *xií* *bait-áo-p-i* ***"híx***
cloth wash-TELIC-IMPF-PROX Q
'Are you going to wash clothes?'

b. *xísi* *ib-áo-p-**óxóí***
3.ANIMAL hit.arrow-TELIC-IMPF-Q
'Did you arrow fish?'

A fourth but fairly uncommon way to signal that an utterance is a question is by using a **different word order** from that used in corresponding declarative sentences. This method is used in a number of European languages, like German, in which the inflected verb is placed at the beginning of the sentence. Compare the declarative sentence in (6a) with its corresponding interrogative sentence in (6b).

(6) German

a. *Der* *Lehrer* ***trink-t*** *das* *Wasser.*
the teacher **drink-3SG** the water
'The teacher is drinking the water.'

b. ***Trink-t*** *der* *Lehrer* *das* *Wasser?*
drink-3SG the teacher the water
'Is the teacher drinking the water?'

English employs a version of this method, except that it is only auxiliary verbs that occur initially, as in (7).

(7) ***Are*** *you planning to go to the beach?*

While this method of signalling polar questions is largely confined to Europe, it is also found in two Austronesian languages, Palauan (Austronesian; Palau, Pacific; Josephs 1975: 410) and Manggarai (Austronesian; Nusa Tenggara, Indonesia; Semiun 1993: 25–6), and one Arawakan language, Warekena (Aikhenvald 1998: 261). In all three languages, polar questions employ verb-initial order.

Another uncommon method for coding a sentence as a polar question is by **the absence of morphemes used in declarative sentences**. In Zayse (Omotic, Afro-Asiatic; Ethiopia), the forms of verbs used in declarative sentences contain a morpheme *-tt(e)-* that is absent from corresponding interrogative forms (e.g. *hamá-tte-ten* 'I will go', *háma-ten* 'will I go?', *hamá-tt-isen* 'she will go', *háma-ysen*

'will she go?'; Hayward 1990b: 307). Three other languages of this sort that are shown on the map are Kabardian (North-West Caucasian; Russia; Colarusso 1992: 125–6), Puquina (isolate; Bolivia; Adelaar 2004: 354), and Dinka (Nilotic; Sudan; Nebel 1948: 58–61).

Burunge (Cushitic, Afro-Asiatic; Tanzania) employs a method for signalling polar questions that is difficult to classify, which might be treated as an instance of coding by the absence of a declarative morpheme, but which is treated here as an instance of interrogative verbal morphology. Burunge has a process whereby the final vowel in a declarative sentence is pronounced as a whispered vowel, rather than voiced. This process does not apply, however, in polar questions, so that the polar question differs from the corresponding declarative sentence only in that the final vowel in the sentence is whispered in the declarative sentence, as in (8a), but voiced in the interrogative sentence, as in (8b).

(8) Burunge (Kießling 1994: 213)

a. *'ugu* *qaymo-d-osi*
2SG.M field-F-3SG.POSS
ha-ga *doosl-id^{a}*
NON3.SUBJ-3SG.F.OBJ cultivate-2SG.IMPF.DECL
'You are cultivating his field.'

b. *'ugu* *qaymo-d-osi*
2SG.M field-F-3SG.POSS
ha-ga *doosl-id**a***
NON3.SUBJ-3SG.F.OBJ cultivate-2SG.IMPF.Q
'Are you cultivating his field?'

One might view this contrast as involving a morpheme in declarative sentences whose realization is the whispering of the final vowel in the sentence, and then analyse interrogative sentences as lacking this morpheme, in which case Burunge would be a fifth instance of a language in which polar questions involve the absence of a morpheme that is present in declarative sentences. On the other hand, one might argue that polar questions involve more phonetic material than the corresponding declarative sentence in so far as a voiced vowel has more phonetic content than a whispered vowel. I have somewhat arbitrarily opted for the latter view in classifying Burunge for this map. A very similar contrast between whispered vowels in the last syllable of declarative sentences and voiced vowels in interrogative sentences is found in Nkore-Kiga (Bantu; Uganda; Taylor 1985: 6). However, because Burunge is verb-final, this contrast always occurs on the verb, and thus can be interpreted as part of the verb morphology. In contrast, Nkore-Kiga is SVO and the implication of Taylor's discussion is that the whispered-voiced contrast occurs on the last syllable of the sentence, regardless of the category of the word. The contrast in Nkore-Kiga is thus somewhat analogous to a clitic that attaches to the last word in a sentence (as in the Fyem example in (3) above), and is treated here as an instance of a sentence-final question particle, though again this is subject to alternative analyses.

The sixth type shown on the map involves the same words, morphemes, and word order as the corresponding declarative sentence, but with a distinct **intonation pattern** as the sole indication that it is a question. An example of such a language is colloquial Italian (Maiden and Robustelli 2000: 147). Languages of this type are proportionally underrepresented on the map: there are more languages that employ only interrogative intonation than the map suggests.

Many if not most languages of the first five types also employ a distinct intonation, though some, such as Imbabura Quechua (Ecuador; Cole 1982: 15), do not. And languages in which intonation is the sole indication of a polar question differ as to the nature of the intonational differences. For example, while many languages employ rising intonation at the end of questions, Sesotho (Bantu; South Africa and Lesotho) employs lowered pitch on the final syllable of the sentence (Paroz 1946: 208).

If there is no evidence of any grammatical device other than intonation being used to indicate a neutral polar question in a language, the language is shown on the map as having interrogative intonation only. In some languages, intonation may be the most common means of indicating a polar question, but if some other method is used a minority of the time, then the language is shown on the map according to that method. For example, in Ocuilteco (Otomian, Oto-Manguean; Mexico), the description by Muntzel (1986: 145) suggests that intonation is the normal means but that occasionally polar questions are marked by a sentence-initial particle; hence this language is shown on the map as a language employing a question particle.

The final type shown on the map is languages in which there is **no formal marking of polar questions**, by any of the means represented by the first six types. Only one language in the sample is of this type, Chalcatongo Mixtec (Oto-Manguean; Mexico). The example in (9) could be either a declarative sentence or an interrogative sentence, with no difference in intonation associated with the two meanings.

(9) Chalcatongo Mixtec (Macaulay 1996: 126)
xakú=ro
laugh=2
'You're laughing. / Are you laughing?'

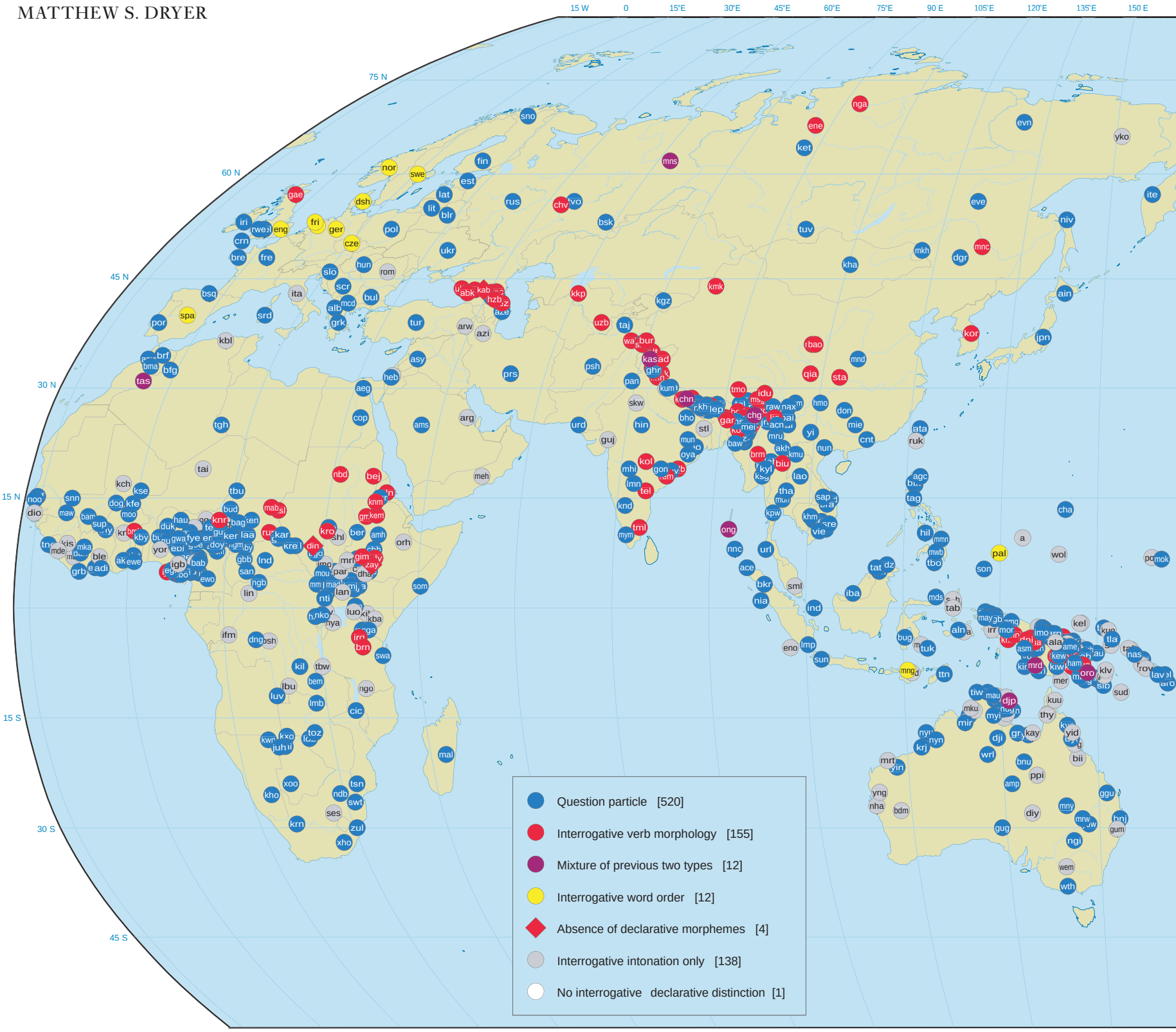
The map restricts attention to ways of signalling unbiased questions, in contrast to leading questions, which signal that the speaker has an expectation as to what the answer to the question will be. See Chapter 92 for further discussion. See also Chapter 140 on question marking in sign languages.

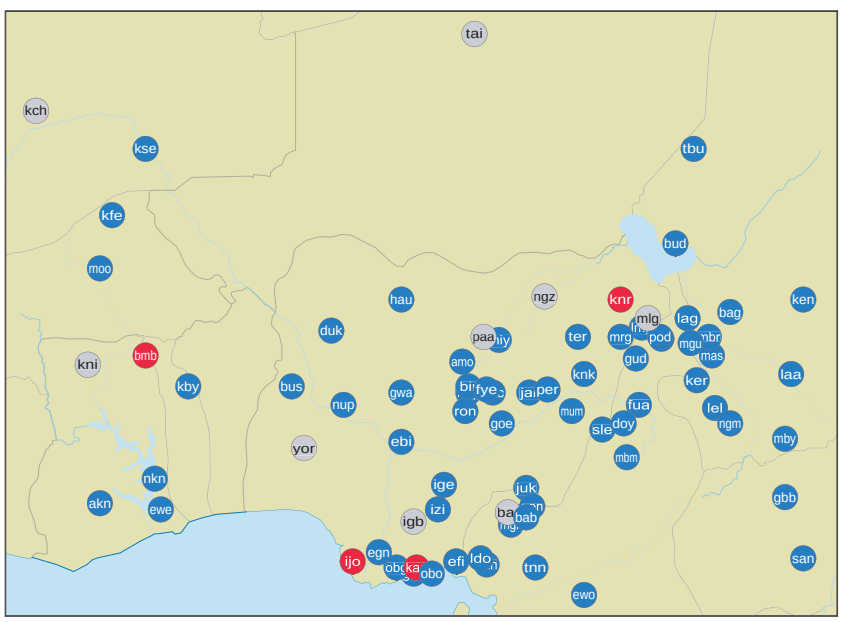
2 Geographical distribution

The map shows that the use of morphemes to signal polar questions, either question particles or interrogative verb morphology, is very widespread. Separate question particles are more common than interrogative verb morphology and are found throughout the world. Interrogative verb morphology is somewhat less widespread: it is not found in Australia or in an area stretching from South-East Asia through the Pacific, except in New Guinea. It is somewhat more frequent in an area in Asia stretching from the Caucasus to Myanmar. Languages employing intonation only are also widespread, though more common in South America, Africa, the Middle East, Australia, and New Guinea than in Europe, much of Asia, and North America. Interrogative word order is largely restricted to western Europe, but the map shows one instance in Indonesia, one in the western Pacific, and one in South America.

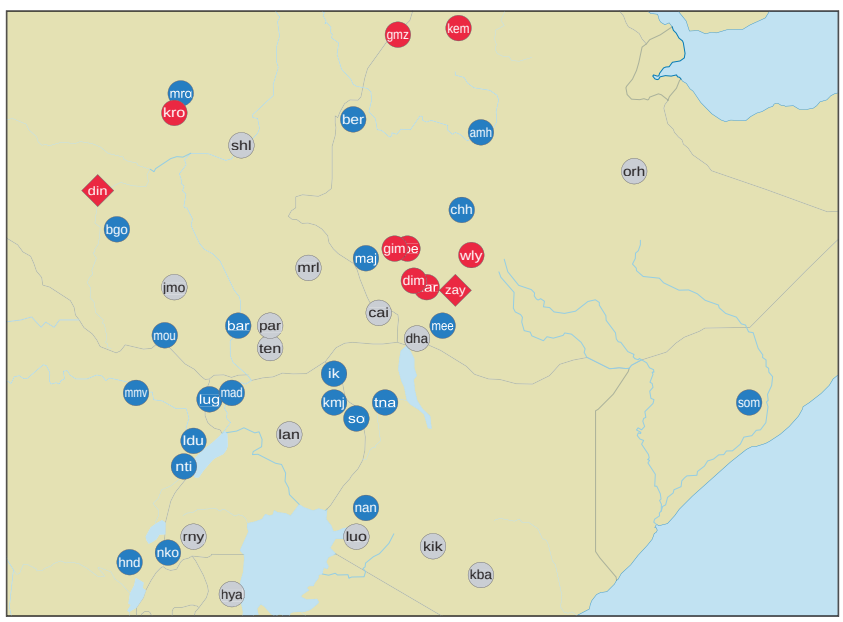
116 Polar Questions

MATTHEW S. DRYER

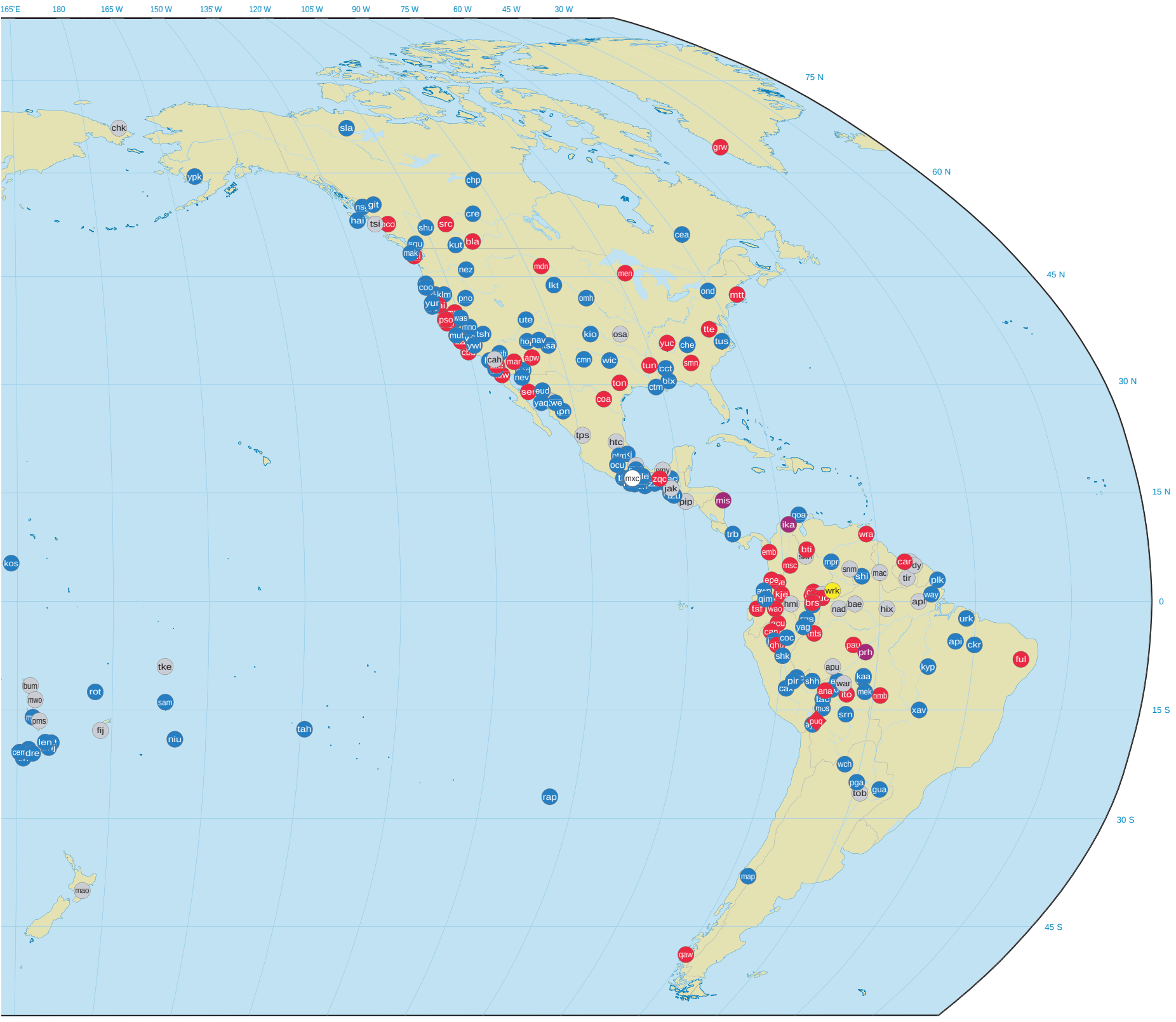


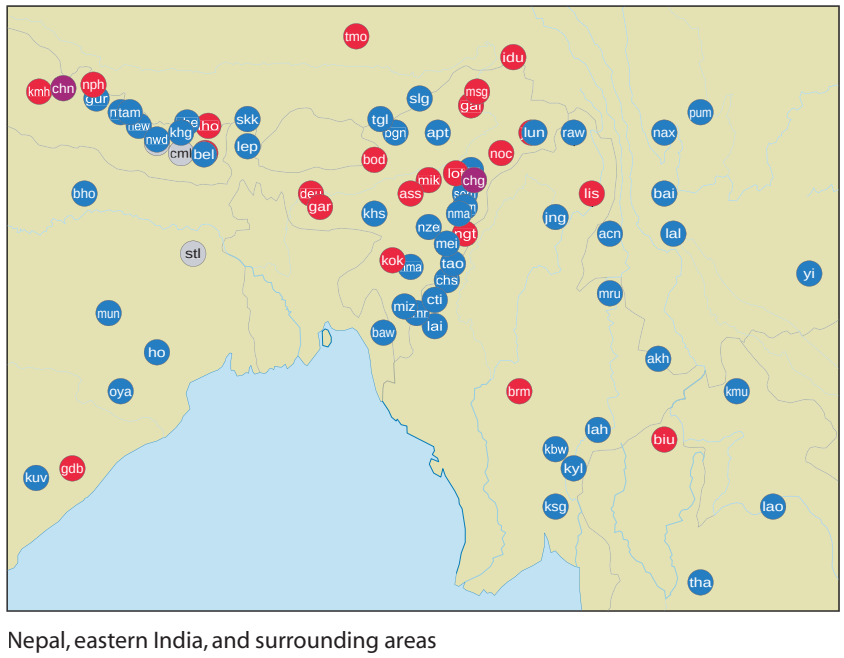


Nigeria and surrounding areas

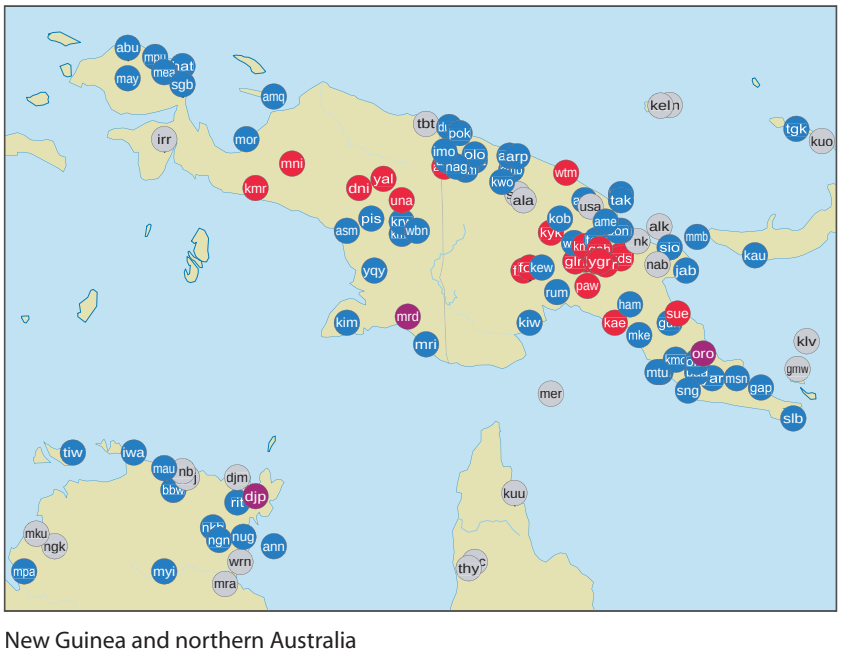


eastern Africa





Nepal, eastern India, and surrounding areas



New Guinea and northern Australia

117 Predicative Possession

LEON STASSEN

1 Defining the construction

This map illustrates the areal distribution of the various encoding options for **predicative possession.** The database for this map thus includes sentences in which ownership of a certain object (the possessed item) is predicated of a possessor, in a way that is illustrated by the English sentence *John has a motorcycle.* It should be noted that the scope of this map is restricted by several conditions. First, the map deals only with the encoding of **predicative** possession, as in *John has a motorcycle*; cases of adnominal possession (as in *John's motorcycle*) are not dealt with. Secondly, within predicative possession it often matters whether the noun phrase that indicates the possessed item is **indefinite** or not. English is a language in which this parameter gives rise to two different encoding options (see 1a-b). For the purposes of this map, only constructions in which the possessed NP has an indefinite reading—i.e. only those constructions that are parallel to (1a)—have been taken into account.

(1) a. *John has a motorcycle.*
b. *This motorcycle is John's.*

Finally, only those expressions have been sampled which encode the concept of **alienable possession**. Thus, the map concentrates on the encoding of the domain of "ownership" in a narrow juridical or ethical sense; it comprises those cases in which the relation between the possessor and the possessed item can be disrupted, transferred, or given up by acts of stealing, borrowing, or selling. As a result of this restriction, the encodings of other forms of possession (such as inalienable possession or temporary possession) are not relevant to this map.

2 Major types of predicative possession

Linguists writing on the typology of predicative (alienable) possession (such as Clark 1978; Seiler 1983; Heine 1997; Stassen 2001) agree that there are at least five relatively frequent and easily identifiable encoding strategies. Among these five strategies, one stands apart in that it encodes the possessive relationship between possessor and possessed item in the form of a **transitive construction**. In this **Have-Possessive**, the possessor NP and the possessed NP function respectively as the subject and the direct object of a 'have'-verb, which, in many cases, can be shown to derive from some verb indicating physical control or handling, such as 'take', 'grasp', 'hold', or 'carry'. The English construction presented in (1a) is an instance of the Have-Possessive. Another example is from West Greenlandic (Eskimo); here the possessed item happens to be incorporated into the 'have'-verb.

(2) West Greenlandic (Fortescue 1984: 171)
angut *taanna* *qimmi-qar-puq*
man that dog-have-3SG.IND
'That man has dogs.'

Opposed to the Have-Possessive, the other four major types employ a strategy which is syntactically intransitive: the possessive construction has the basic form of an **existential sentence**. Thus, all three of these types feature a one-place predicate with a locational or existential meaning; its usual translation can be something like 'to be at', 'to be there', 'to exist'. The difference between these types lies in the encoding of the possessor NP and the possessed NP.

In the **Oblique Possessive**, the possessed NP functions as the grammatical subject of the 'exist'-predicate, while the possessor NP is constructed in some oblique form. The Oblique Possessive has two subtypes. In one, the oblique marking on the possessor NP has as its basic meaning the specification of a locational relation. Depending on the particular type of locational relation selected, it would be possible to further subcategorize this type into Locative Possessive (with the possessor NP being marked by some item meaning 'at', 'on', or 'in') and Dative Possessive (with a marker 'to' or 'for' on the possessor NP). On this map, however, these differences are ignored, and all instances of locational marking have been brought together under the heading of the **Locational Possessive**. An example of the Locational Possessive comes from Written Mongolian.

(3) Written Mongolian (Poppe 1954: 147)
na-dur *morin* *bui*
1SG-at horse be.3SG.PRES
'I have a horse.' (lit. 'At me is a horse.')

The second subtype of the Oblique Possessive is the **Genitive Possessive**. Here the possessor NP is marked by an item which typically does not have a locational interpretation; moreover, the possessor NP is commonly (though not necessarily) constructed as an adnominal modifier to the possessed NP. The Genitive Possessive is illustrated by an example from Avar (Daghestanian; Caucasus).

(4) Avar (Kalinina 1993: 97)
dir *mašina* *b-ugo*
1SG.GEN car III-be.PRES
'I have a car.'

The **Topic Possessive** shares with the Locational and the Genitive Possessive the characteristic that the possessed NP is construed as the grammatical subject of the existential predicate. The distinguishing feature of the Topic Possessive lies in the encoding of the possessor NP, which is construed as the **topic** of the sentence. As such, the possessor NP indicates the "setting" or "background" of the sentence, that is, the discourse frame which restricts the truth value of the sentence that follows it. Its function can thus be paraphrased by English phrases such as *given X*, *with regard to X*, *speaking about X*, *as far as X is concerned*, and the like. An example is from Tondano (Austronesian; northern Sulawesi).

(5) Tondano (Sneddon 1975: 175)
si *tuama* *si* *wewean* *wale* *rua*
ANIM.SG man TOP exist house two
'The man has two houses.' (lit. 'As far as the man is concerned, there are two houses.')

Like the other intransitive possessive types, the **Conjunctional Possessive** contains an existential predicate. In other respects, however, the Conjunctional Possessive contrasts with both the Oblique Possessive and the Topic Possessive. For a start, the Conjunctional Possessive construes the possessor NP as the grammatical subject. An even more conspicuous feature is the encoding of the possessed NP. In the Conjunctional Possessive this NP is accompanied by, and usually in construction with, a marker which can be analysed neither as a locational item nor as an indicator of topic. Closer inspection reveals that this marker in all cases originates from an item which is, or at least has been, employed as a means of indicating **simultaneity between clauses**. Thus, we find markers which have their origin in a sentential adverb meaning 'also' or 'too', or in a subordinating conjunction 'when/while', or in a coordinating particle 'and'. A prominent option within the Conjunctional Possessive is the use of the comitative marker 'with' on the possessed NP, which is why the type is often referred to in the literature as the **WITH-Possessive**. It can be argued, however, that languages which employ this comitative marker on possessed NPs also use this marker as a means to coordinate noun phrases (see Stassen 2000), so that this WITH-strategy can be seen as a special case of a more general conjunctional encoding format. Examples of the Conjunctional Possessive are from Daga (Dagan, Trans-New Guinea; southeastern

Papua New Guinea) and Sango (Adamawa-Ubangi; Central African Republic).

(6) Daga (Murane 1974: 303)
orup da agoe den
man one slave with/too
'A man had a slave.'

(7) Sango (Samarin 1967: 95)
lo eke na bongo
3SG be and/with garment
'She has a garment.'

In accordance with the above discussion, the following values are shown on the map:

1. Locational Possessive	48
2. Genitive Possessive	22
3. Topic Possessive	48
4. Conjunctional Possessive	59
5. Have-Possessive	63
total	240

3 Grammaticalized possessive constructions

A number of languages exhibit constructions which cannot be classified straightforwardly in terms of any of the five basic types. Closer inspection reveals that these cases can be rated as the results of several grammaticalization processes. First, we can note a phenomenon that may be called **Transitivization** or HAVE**-Drift**, as it consists in a process of drifting from one of the other basic types towards a Have-Possessive. Cases of HAVE-Drift from an erstwhile Conjunctional Possessive commonly involve the cliticization or incorporation of the conjunctional marker into the existential predicate; the newly formed predicate then acts as a transitive verb. An example is from Luganda (Bantu; Uganda).

(8) Luganda (Ashton et al. 1954: 234)
o-li-na ekitabo
2SG-be-with book
'You have a book.'

HAVE-Drift from Topic Possessives commonly involves the reanalysis of the existential 'be'-item as a transitive verb, and the reanalysis of the possessor NP and possessed NP as the subject and direct object of that verb, respectively. The process is helped along by the fact that, in the typical case, this 'be'-item occupied the canonical position of transitive verbs in the original possessive construction. That the process is gradual and involves various intermediate stages can be seen from sentences from Luiseño (Uto-Aztecan; southern California). Here the reanalysis of the erstwhile topic into subject seems to be under way, but the construction is not yet unequivocally transitive.

(9) Luiseño (Steele 1977: 114, 122)

a. *noo-p no-toonav qala*
1SG-TOP my-basket be.INAN.PRES
'I have a basket.'

b. *noo-n no -toonav qala*
1SG-SUBJ my-basket be.INAN.PRES/have
'I have a basket.'

Instances of HAVE-Drift from Locational Possessives are not very frequent. Comrie (1989: 219–25) reports a case from Maltese (Semitic; Malta) involving an intermediate stage in which the possessor NP is topicalized. A similar process must have taken place in the Celtic language Breton (see Press 1986: 139). On the map, instances of HAVE-Drift have not been marked separately; they are subsumed under their respective source types.

A second instance of grammaticalization of predicative possessive structures might be called **Adjectivalization.** In some linguistic areas, we find possessive constructions in which the possessed NP is construed as the predicate (or part of the predicate) and treated in the same way as predicative adjectives are treated. Thus, depending on whether predicative adjectives are "nouny" or "verby" (Wetzer 1996, Stassen 1997), the possessed noun phrase shows up as (part of) the complement of the copula, or as (the lexical core of) a predicative verb. Examples are from Tiwi (Bathurst Island, northern Australia), Kanuri (Saharan; northern Nigeria), Guajajara (Tupian; Maranhão State, Brazil), and Kolyma Yukaghir (isolate; northeastern Siberia).

(10) Tiwi (Osborne 1974: 60)
ŋawa mantani ʈəraka
our friend wallaby
'Our friend has a wallaby.'

(11) Kanuri (Cyffer 1974: 122)
kâm kúrà-tá kúŋŋɔ́nà-nzɔ́-wà (gànyí)
man big-the money-his-ADJ/with (NEG.COP)
'The big man has (no) money.'

(12) Guajajara (Bendor-Samuel 1972: 162)
i-mukaw
3SG-gun
'He has a gun.'

(13) Kolyma Yukaghir (Jochelson 1905: 405)
Met āče-n·-je
I reindeer-with-1SG.PRES.INDEF.INTR
'I have (a) reindeer.'

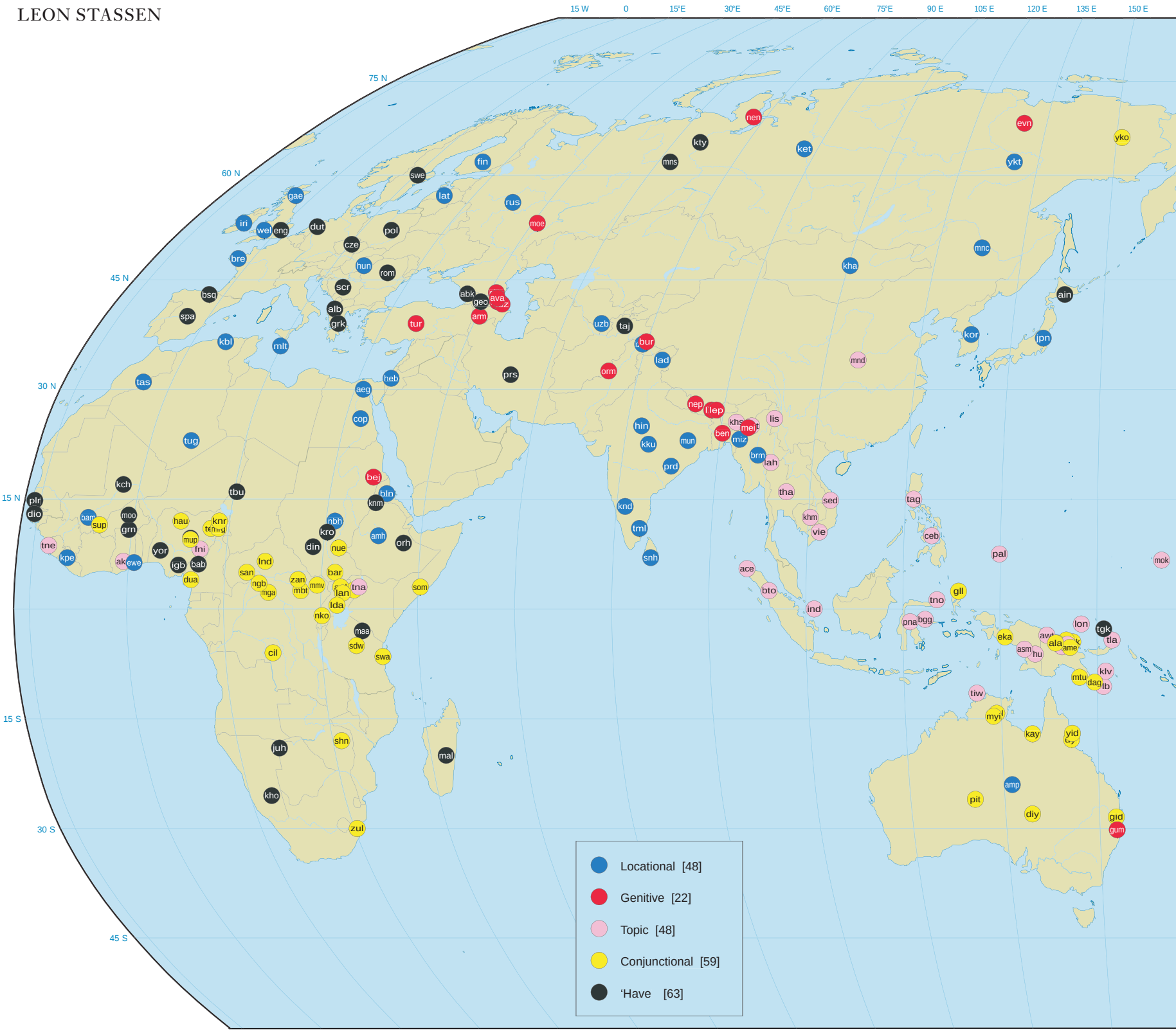
Cases like these are probably best viewed as the result of a grammaticalization process by which the possessed noun phrase (together with its marker, if it has one) is gradually reanalysed as the predicate of the construction. Depending on whether the possessed noun phrase bears a marker or not, the source of such products of adjectivalization can be traced back to a Conjunctional Possessive or a Topic Possessive. Therefore, cases of adjectivalization are not represented separately on the map, but are coded in accordance with their source type.

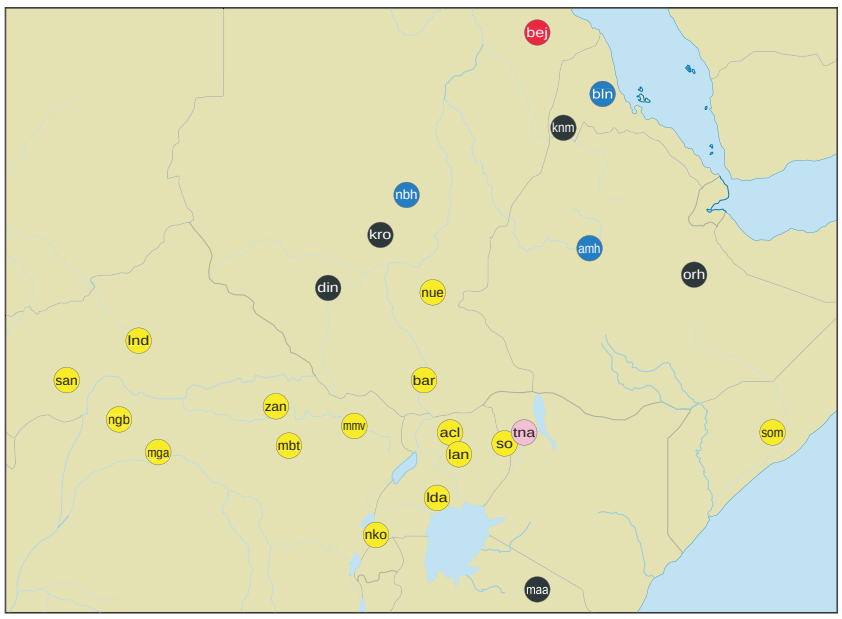
4 Geographical distribution

As the map demonstrates, the distribution of the various types of predicative possession shows considerable areal effects. Eurasia and North Africa (with the exception of the languages of western Europe) are almost exclusively the domain of the Oblique Possessive. Further areas where this encoding option is found are Polynesia and the northern part of South America. For the Topic Possessive, the major area is East/South-East Asia (including Indonesia, the Philippines, and parts of New Guinea); the option is also found in many language groups of the Americas, and in some areas in western and northeastern Africa. The Conjunctional Possessive is dominant in sub-Saharan Africa and in Australia and New Guinea, but the option also presents itself all over the Americas. Finally, a core area for the Have-Possessive is western and central Europe; however, we also find quite a few instances of this option in Africa and in the Americas.

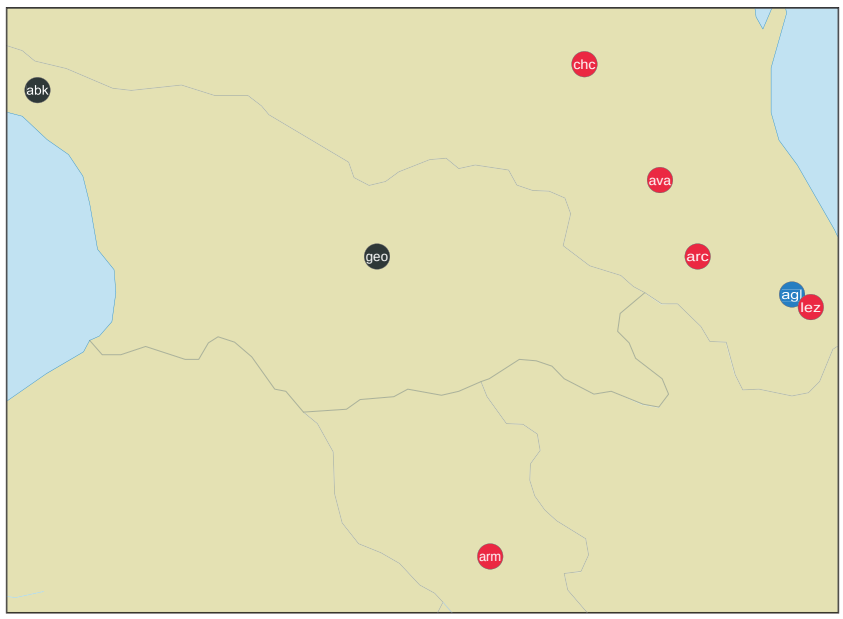
117 Predicative Possession

LEON STASSEN

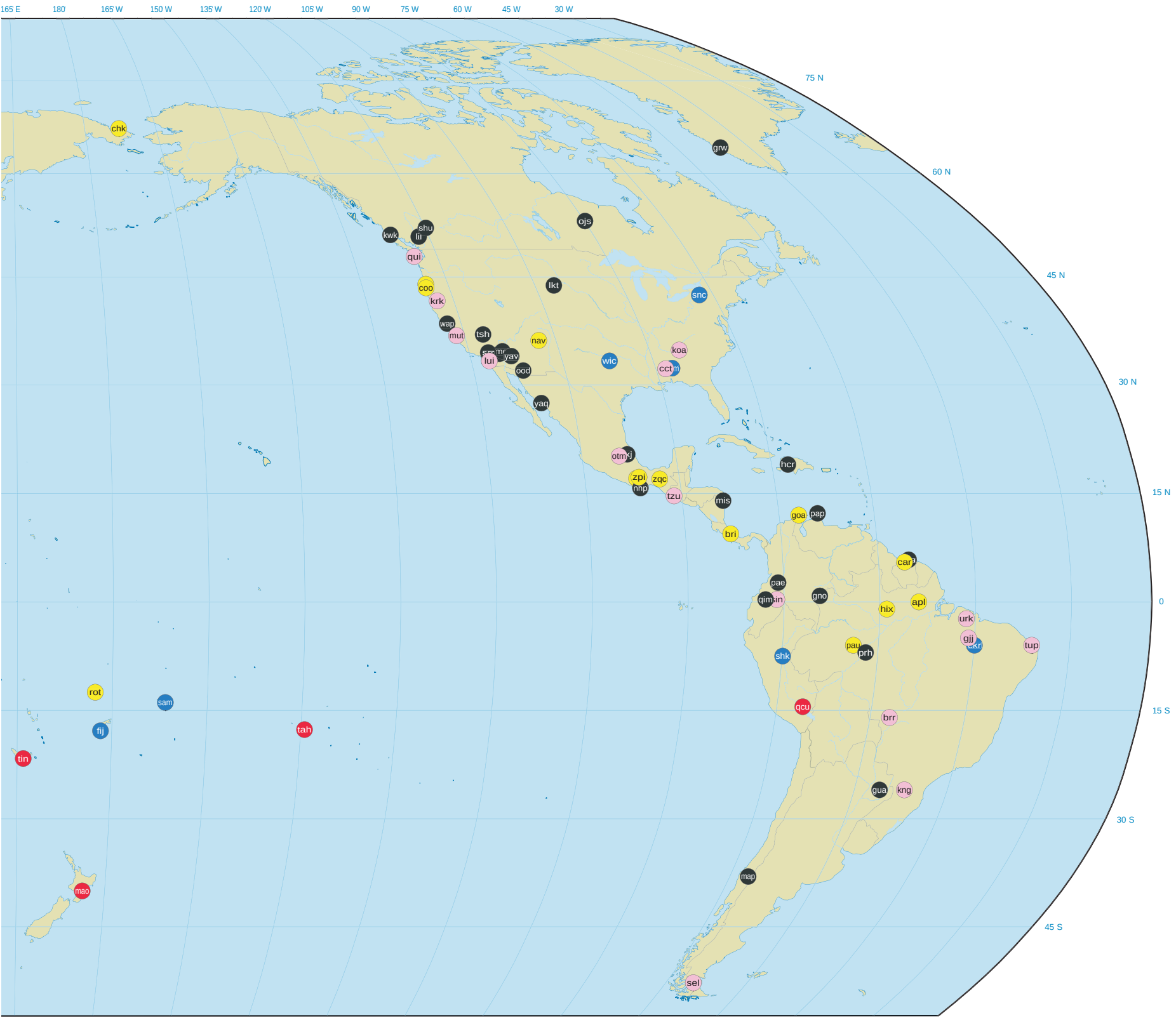


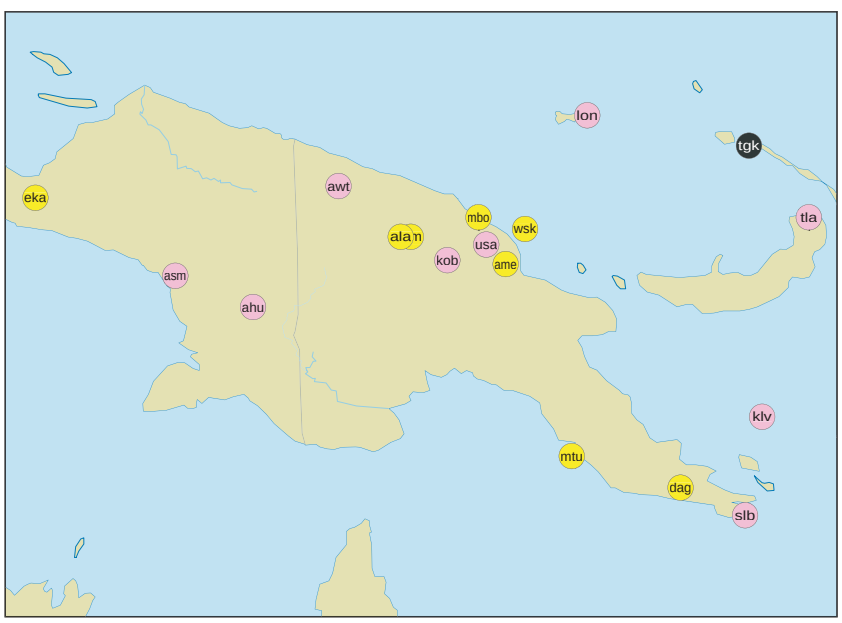


eastern Africa

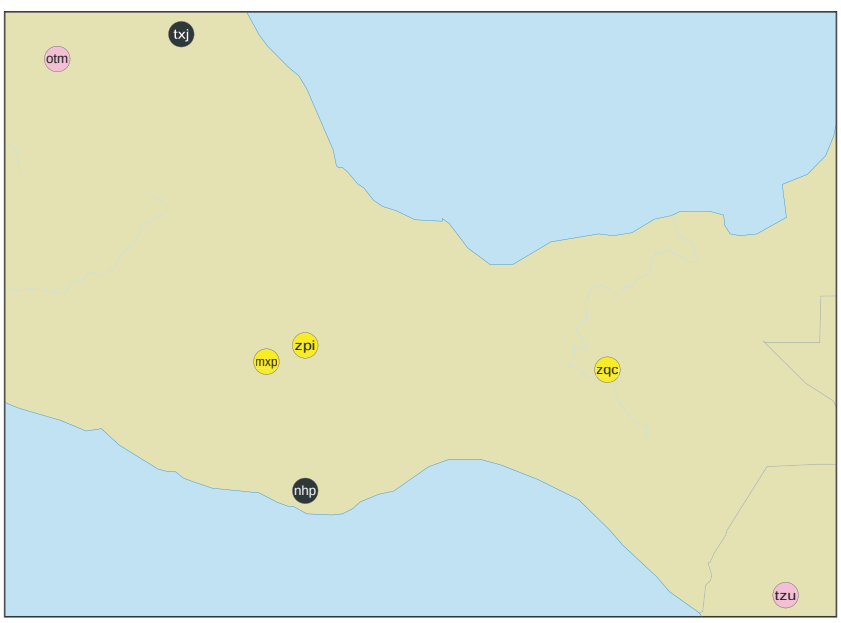


Caucasus





New Guinea



southern Mexico

118 Predicative Adjectives

LEON STASSEN

1 Defining the values

This map shows the distribution of the various options in the encoding of predicative adjectives, i.e. items which predicate a property of a subject. The basic distinction is between those languages in which predicative adjectives are encoded in a way that is parallel to predicative verbs, and those languages in which the encoding of predicative adjectives and of verbs is different. An example of this latter, **nonverbal** encoding of predicative adjectives is English. As shown in the sentences in (1), English shows third-person agreement in the present tense of its predicative verbs, but this option is not available for predicative adjectives: **John tall-s* is not acceptable in the language. On the other hand, Bororo (Macro-Ge; Mato Grosso, Brazil) offers an instance of **verbal** encoding of predicative adjectives; as the sentences in (2) demonstrate, the encoding of predicative verbs and of predicative "property words" is identical.

(1) a. *John sleep-s.*
b. *John is tall.*

(2) Bororo (Crowell 1979: 26, 50)
a. *i-mago-re*
1SG-speak-NEUTRAL
'I speak/spoke.'
b. *i-kuri-re*
1SG-tall-NEUTRAL
'I am/was tall.'

In addition to the two basic options of verbal and nonverbal encoding, the map shows a third encoding type for predicative adjectives. In this **mixed** type, both a verbal and a nonverbal encoding of predicative adjectives is available. An example is from Luo (Nilotic; Kenya). Sentence (3b) shows that, in this language, the item *be'r* 'good' can have verbal encoding, while sentence (3c) indicates that it may also have nonverbal encoding.

(3) Luo (Tucker and Bryan 1966: 425, 432)
a. *à-lwóŋó*
1SG-call.NONPERF
'I am calling.'
b. *à-bê·r*
1SG-good.NONPERF
'I am good.'
c. *án* *má-bê·r*
1SG.EMPH NMLZ-good
'I am good.'

In accordance with the above, the following values are shown on the map:

1.	Predicative adjectives have verbal encoding	151
2.	Predicative adjectives have nonverbal encoding	132
3.	Predicative adjectives have mixed encoding	103
	total	386

2 Criteria for verbal and nonverbal encoding

In order to be able to decide whether a given predicative adjective construction is a case of verbal or nonverbal encoding, we need a set of cross-linguistically applicable criteria. For this map, the following three criteria have been employed (see Wetzer 1996; Stassen 1997). The first of these is the Agreement Criterion:

(4) **The Agreement Criterion**
If a language has (person/number/gender) agreement on predicative verbs, then predicative adjectives in that language will be rated as verbal if they show this agreement marking as well. If they do not, they will be rated as nonverbal.

The application of this criterion to English and Bororo enables us to decide that English has nonverbal adjectives, whereas Bororo has verbal adjectives (see 1–2). Another case in which this criterion is decisive is presented by Tiwi (Australian; Bathurst Island). As the sentences in (5) demonstrate, predicative adjective encoding in Tiwi must be rated as nonverbal, since the agreement marking that is required by verbs is lacking on predicative adjectives.

(5) Tiwi (Osborne 1974: 70, 60)
a. *ji-pauliɣi*
3SG M.PST-fall
'He fell.'
b. *ţuŋkwaltiriŋa* *pumpuka*
stringy.bark good
'The stringy bark is good.'

A second criterion which—alone, or in tandem with the Agreement Criterion—can decide on verbal or nonverbal status of predicative adjectives is the Copula Criterion.

(6) **The Copula Criterion**
If predicative adjectives are marked by the presence of a supportive item (a copula), then their encoding must be rated as nonverbal.

Application of this criterion yields—in addition to the Agreement Criterion—the conclusion that English predicative adjectives are nonverbal, since they require the copular item *be*. A case in which the Copula Criterion is the sole decisive factor is presented by Irish. In this language, predicative verbs do not show any agreement marking, so that the Agreement Criterion is not applicable here. Nonetheless, we can decide upon nonverbal status for Irish predicative adjectives, by virtue of the Copula Criterion.

(7) Irish (Greene 1966: 46, 43)
a. *téann* *Sean*
go.PST Sean
'Sean went.'
b. *is* *breoite* *é*
COP.PRES ill he
'He is ill.'

For languages that lack both copulas and agreement marking on verbs, the two criteria presented above are not applicable. For this case, we need an additional third criterion, the Negation Criterion.

(8) **The Negation Criterion**
If predicative verbs and adjectives show different negation, then the encoding of predicative adjectives must be rated as nonverbal.

The application of this criterion can be illustrated by Gumbaynggir (Pama-Nyungan; New South Wales, Australia) and Tagalog (Austronesian; Philippines). Both languages lack copulas and agreement on verbs. However, Gumbaynggir has different negation strategies for verbs and adjectives (Eades 1979: 332), while the negation strategy in Tagalog is the same for the two predicate categories (see 9a-b). Hence we determine that predicative adjectives are verbal in Tagalog, whereas they are nonverbal in Gumbaynggir.

(9) Tagalog (Schachter and Otanes 1972: 518)
a. *hindi* *dumating* *ang* *bus*
NEG come TOP bus
'The bus did not come.'

b. *hindi* *mura* *ang* *karne*
NEG cheap TOP meat
'The meat is not cheap.'

By applying this set of three criteria it is possible to decide on the verbal or nonverbal status of any case of predicative adjective encoding in a language-independent fashion. It should be stipulated here that the three criteria are ranked as to their application. That is, the first criterion to be applied is the Agreement Criterion. For those languages to which this criterion cannot be applied (because of a lack of agreement on verbs), the Copula Criterion comes into play. Only in cases where the Copula Criterion still does not provide a decision (because the language in question has a zero-copula for predicate nominals) is the Negation Criterion brought in as a final arbiter on the status of predicative property items.

3 Mixed encoding

Mixed encoding of predicative adjectives comes in two types. The example from Luo, presented in (3), illustrates a case of **switching**: the same lexical item (in this case, the item *be'r* 'good') can "switch" between verbal and nonverbal encoding. In the typical instances of switching, the double option of encoding affects only a subset of the predicative adjectives, namely, those for which both a "permanent/inherent" and a "temporary/accidental" interpretation is possible. In such a case, the nonverbal encoding commonly indicates that the property is inherent and/or permanently applicable to the subject. Thus, sentence (3c) (nonverbal) must be interpreted as "I am a good person", whereas sentence (3b) (verbal) indicates that the property of being good is only temporary and/or accidental. A second example of this switching phenomenon is presented by Maori (Polynesian; New Zealand). As the examples in (10) illustrate, the item *pai* 'good' can be encoded either verbally (sentence 10b, on a par with sentence 10a) or nonverbally. In this latter case, the predicative adjective takes a zero-copula, as do predicate nominals in Maori (see sentences 10c and 10d).

(10) Maori (Biggs 1969: 27, 17, 24, 75)

a. *ka* *oma* *te* *kootiro*
INCEP run ART.DEF girl
'The girl runs.'

b. *ka* *pai* *te* *whare* *nei*
INCEP good ART.DEF house this
'This house is good.'

c. *he* *pai* *te* *koorero*
ART.INDEF good ART.DEF talk
'The talk is good.' (lit. 'The talk (is) a good one.')

d. *he* *kiwi* *teera* *manu*
ART.INDEF kiwi this bird
'This bird is a kiwi.'

The second form in which mixed encoding of predicative adjectives manifests itself is that of **split encoding**. In this case, all predicative adjectives have only a single encoding option, but the set of property words is split into a subset with verbal encoding and a subset with nonverbal encoding. An example of this situation can be found in Rama (Chibchan; Nicaragua). As can be seen from the sentences in (11), the item *angaling* 'hungry' gets verbal encoding, whereas the item *mliima* 'good' is encoded nonverbally (by virtue of the Agreement Criterion).

(11) Rama (Colette Grinevald, p. c.)

a. *m-upluul-i*
2-dream-PRES
'You are dreaming.'

b. *nsut* *tiiskibadut* *s-angaling-i*
1PL children 1PL-hungry-PRES
'We children are hungry.'

c. *ning* *suurak* *mliima*
this pineapple good
'This pineapple is good.'

As was the case with switching, there are indications that split encoding of predicative adjectives is often governed by considerations of permanency. Although in many languages the situation is far from clear, there seems to be a tendency to apply the verbal strategy with less permanent properties such as 'hungry', 'ill', or 'sad', and the nonverbal strategy with time-stable properties like 'female' or 'golden' (see Stassen 1997: 164–79).

For the purposes of this map, the differentiation between the two types of mixed encoding has been ignored. Thus, a case of mixed encoding on the map may either stand for switching or for split encoding.

4 Geographical distribution

The map demonstrates that the distinction between verbal and nonverbal encoding of predicative adjectives has clear areal features. Nonverbal encoding appears to be concentrated in two large linguistic areas. The first of these macro-areas comprises all the languages of Europe (with the notable exception of the North-West Caucasian languages), Central Asia, and Siberia (with the exception of Yukaghir), India, the Middle East, and northern Africa.

A second macro-area of nonverbal encoding is formed by (most of) the languages of Australia and New Guinea. Especially along the coastlines of both islands there are a number of counterexamples. Also, there is a certain degree of mixed encoding here, but the central highlands of New Guinea and the bulk of the Australian mainland contain almost uniform nonverbal encoding.

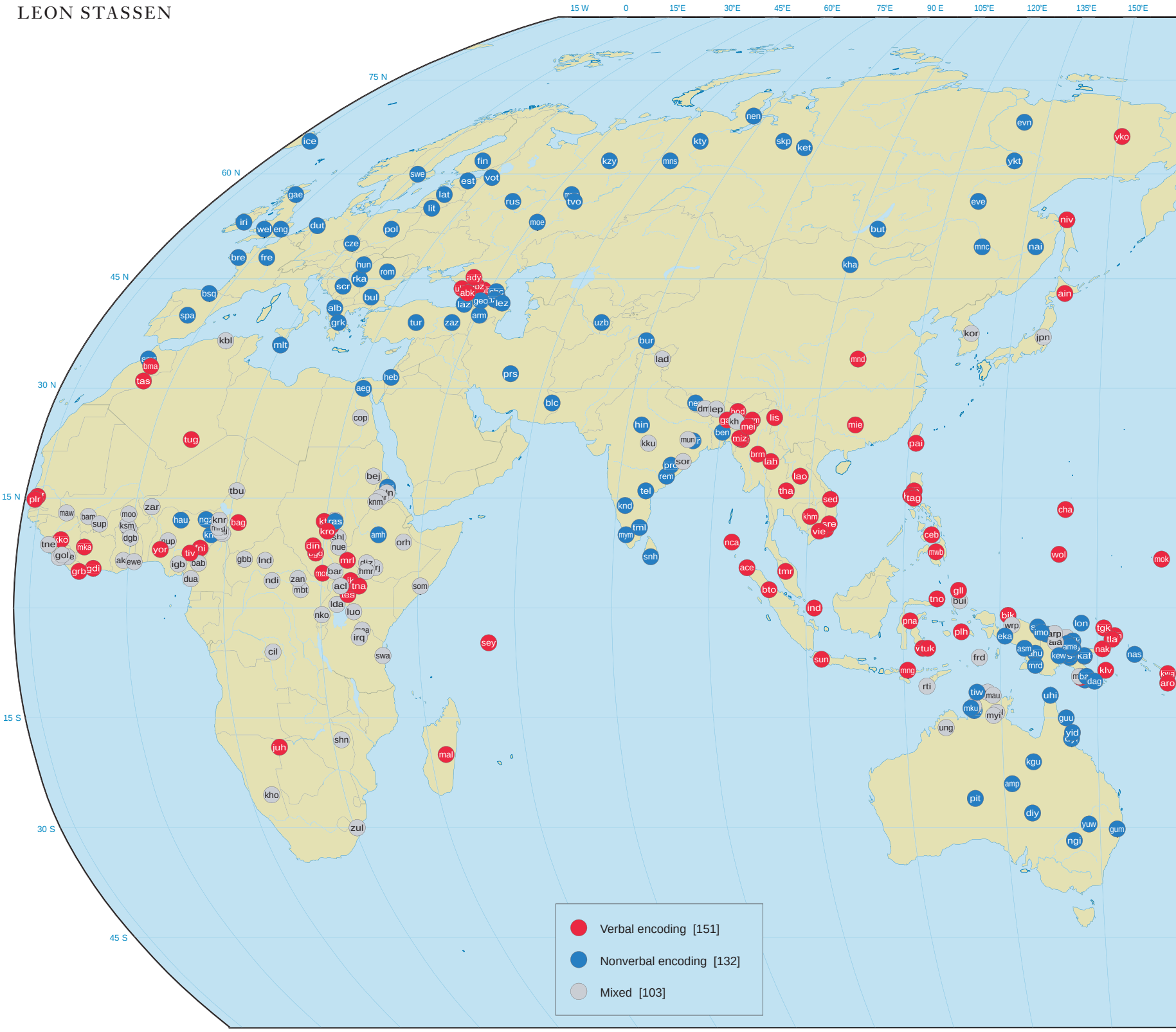
Outside of these two main areas, nonverbal encoding does not seem to be particularly strong. Perhaps the best case for a third nonverbal area can be made for the languages of the southern part of Central America and the eastern part of South America, but there is considerable diversity in this area.

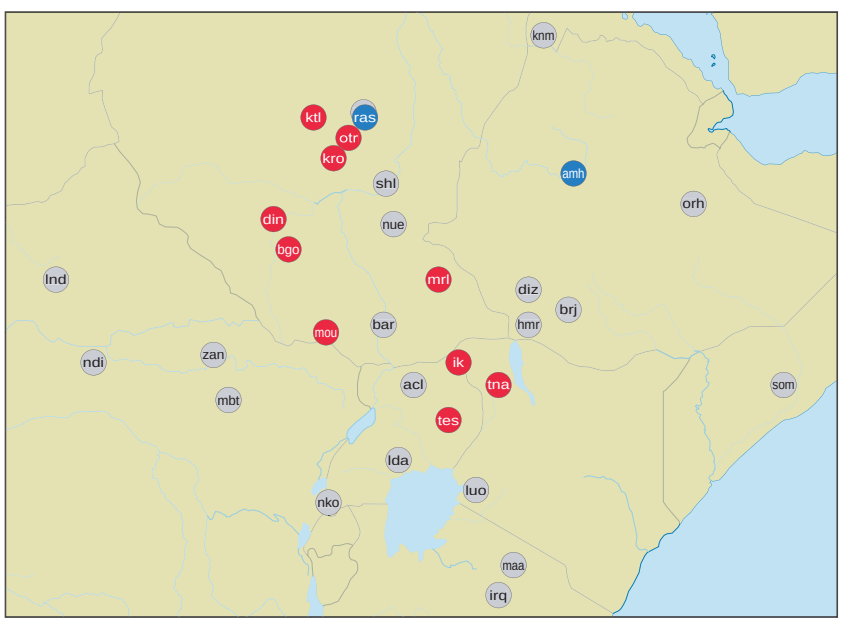
Large unbroken concentrations of verbal encoding are encountered in Africa, Asia, and the Americas. Practically all the languages of Africa in and south of the Sahara exhibit (some degree of) verbal encoding of predicative adjectives. A second area of verbal encoding comprises East and South-East Asia, and the islands of Indonesia, Melanesia, and Polynesia. Furthermore, verbal encoding is the near-exclusive option in North America, and it is prominent in the south and the east of South America as well.

As could be expected, mixed encoding tends to appear in those areas where verbal and nonverbal encoding meet. The most conspicuous area in this respect is sub-Saharan Africa, but mixed encoding can also be found in other borderline regions, such as New Guinea, Central America, and eastern India/Myanmar.

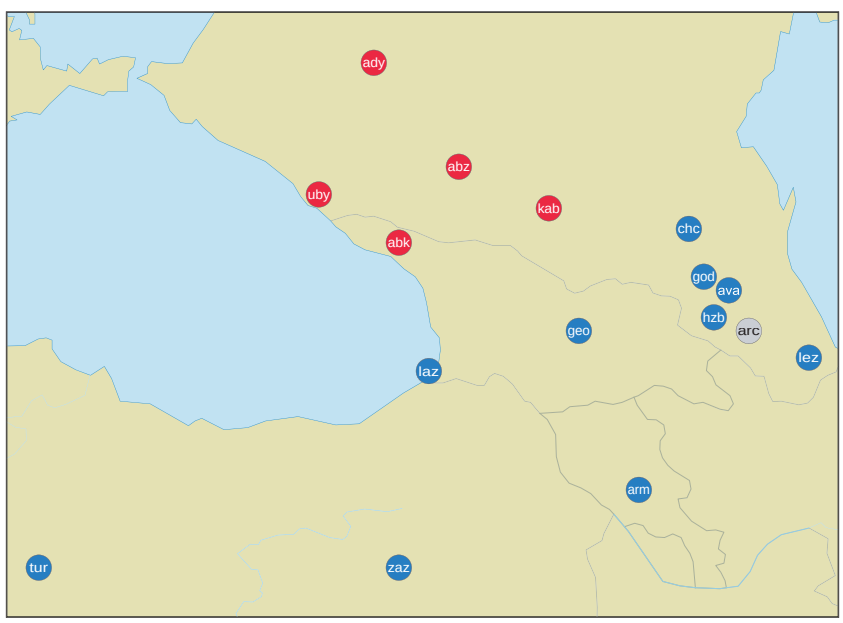
118 Predicative Adjectives

LEON STASSEN

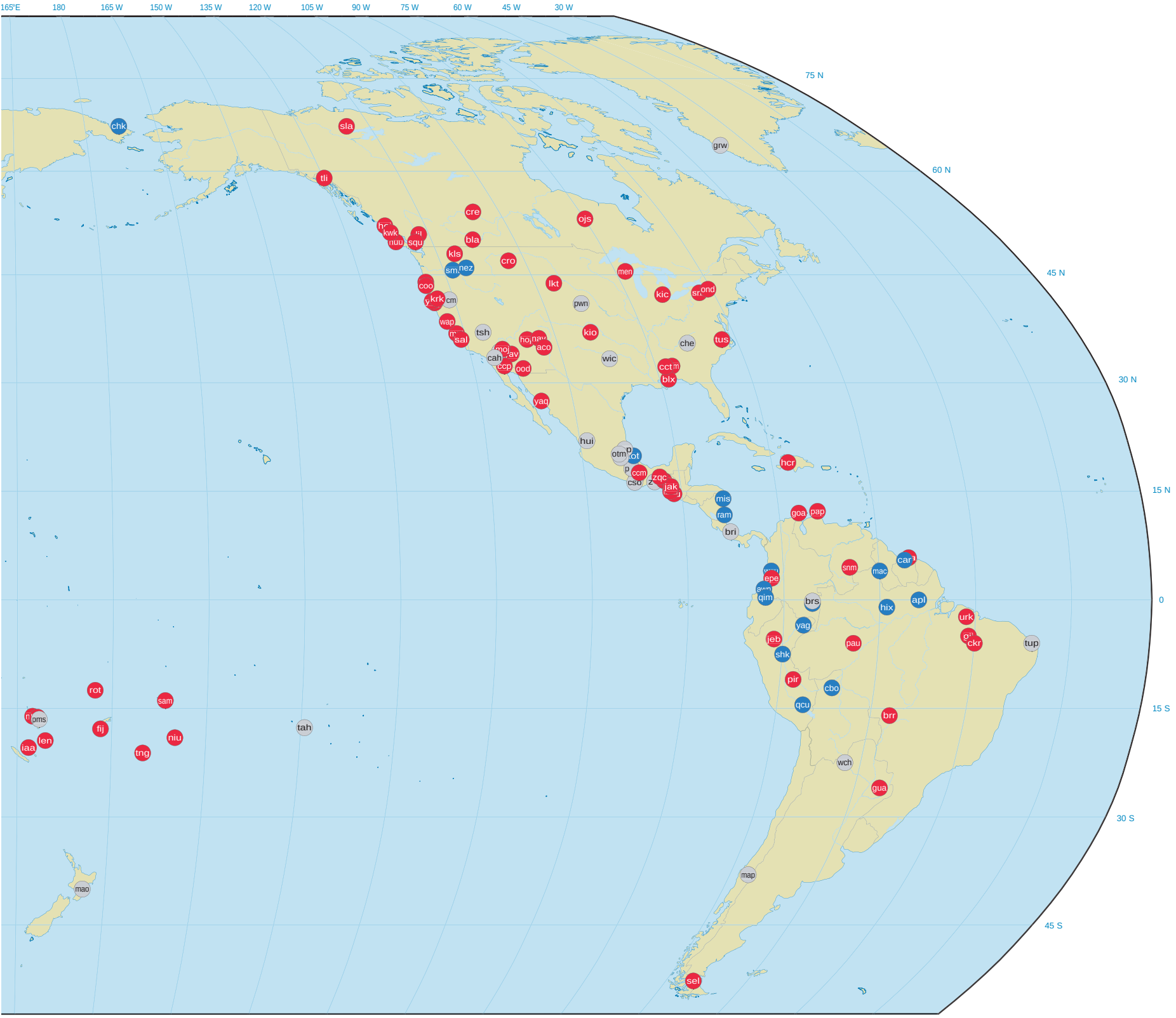


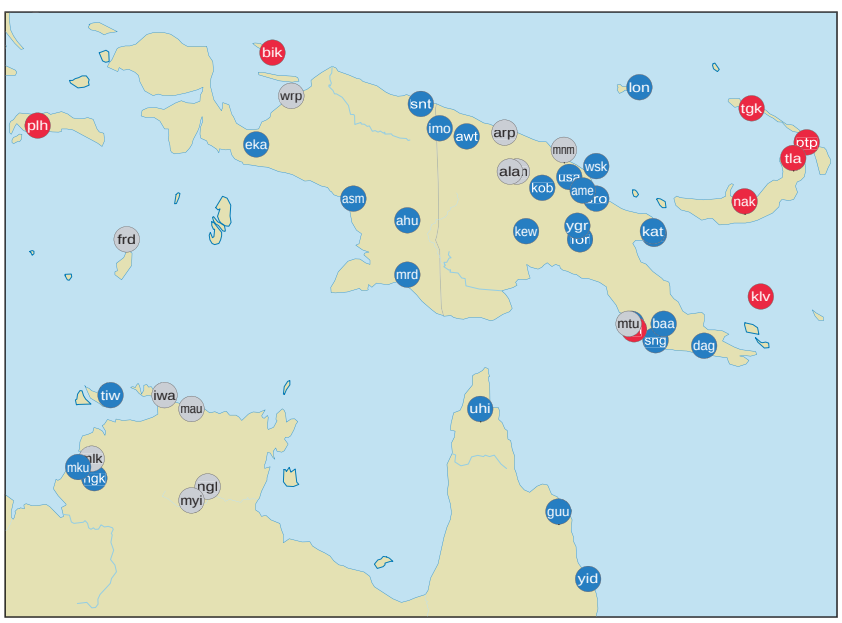


eastern Africa

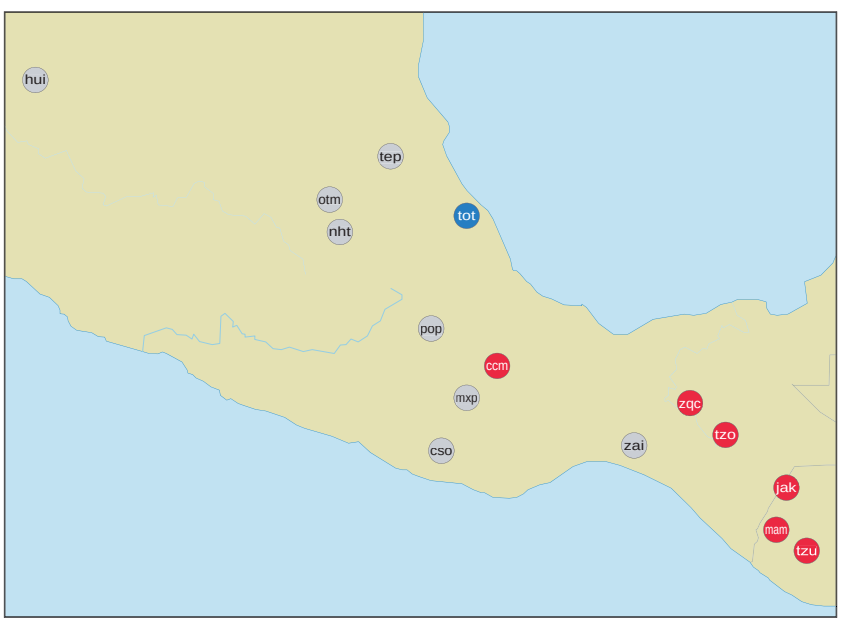


Caucasus and Black Sea area





New Guinea and northern Australia



southern Mexico

119 Nominal and Locational Predication

LEON STASSEN

1 Defining the values

This map shows the possible relationships between the encoding of nominal and locational predicates. In particular, the map shows for each language whether nominal predications (such as *John is a tailor*) and locational predications (such as *John is in Paris*) can or cannot be encoded by the same strategy. In the terminology of Stassen (1997), a language is called a **share-language** if the encoding strategy for locational predications is (or can be) used for nominal predications, and a **split-language** if the encoding strategies for the two constructions must be different.

An obvious example of a share-language is English. As the above example sentences demonstrate, this language can use the lexical item *be* both as a nominal copula and as a locational support verb. In contrast to this, Mandarin is a split-language, as the copula and the locational verb are not the same.

(1) Mandarin (Li and Thompson 1977: 422; Li and Thompson 1981: 365)

a. *nèi-ge rén shì xuéshēng*
that-CLF person COP student
'That man is a student.'

b. *Lǐsì zài hǎi-biān*
Lisi be.at ocean-side
'Lisi is by the ocean.'

The following values are shown on the map:

1.	Split (i.e. different) encoding of nominal and locational predication	269
2.	Shared (i.e. identical) encoding of nominal and locational predication	117
	total	386

2 Variation in split-languages

As was shown in example (1), Mandarin is a split-language by virtue of the difference between the lexical items involved in nominal and locational predication. This type of split encoding is rather common. Spanish, with its difference between the copula *ser* and the locational verb *estar*, is another well-known case in point, as is Irish.

(2) Spanish (Max Kerkhof, p. c.)

a. *Julia es enfermera*
Julia COP.3SG.PRES nurse
'Julia is a nurse.'

b. *Julia está en Barcelona*
Julia be.3SG.PRES in Barcelona
'Julia is in Barcelona.'

(3) Irish (Greene 1966: 40, 43)

a. *is múinteoir é*
COP teacher he
'He is a teacher.'

b. *tá sé sa tseomra*
be.PRES he in.the room
'He is in the room.'

However, this "lexical" form of split encoding is not the only way in which a language can achieve split status. A second, also fairly frequent type of split encoding involves a contrast between a full supporting verb for locational predication and the absence of any overt linking item (a "zero copula") for nominal predication. Examples of split-languages in which this situation holds are Mokilese (Oceanic; Micronesia) and Waskia (Madang; Papua New Guinea).

(4) Mokilese (Harrison 1976: 142, 209)

a. *John johnpadahk-men*
John teacher-INDEF
'John is a teacher.'

b. *ih mine Hawaii*
he be Hawaii
'He is in Hawaii.'

(5) Waskia (Ross and Natu Paol 1978: 11, 12)

a. *aga bawa taleng-duap*
my brother police-man
'My brother is a policeman.'

b. *kadi mu kawam se bage-so*
man ART house in stay-3SG.PRES
'The man is in the house.'

Finally, a third variant of split encoding is based on the difference between a full support verb for locative predicates and a verbal encoding for nominal predicates. Since there are not that many languages in which predicate nominals are treated as verbs anyway, it will be clear that this variant of split encoding will be less frequent than the other two. An example of this encoding option is the Philippine language Kapampangan: as is shown by sentences (6a-b), predicate nouns in this language have the same morpho-syntactic properties as predicate verbs.

(6) Kapampangan (Mirikitani 1972: 137, 44, 72)

a. *tinerak ya ing anak ku*
dance 3SG ART child my
'My child danced.'

b. *mestro ya ing lalaki*
teacher 3SG ART boy
'The boy is a teacher.'

c. *ati ya ing lalaki king eskwela*
be 3SG ART boy at school
'The boy is in school.'

For the purposes of this map, the three possible forms which split encoding can take have been ignored. Thus, a language is rated as a member of type one if there is split encoding of any sort, regardless of whether this involves a lexical contrast, a zero-verb contrast, or a contrast between verbal and nonverbal encoding.

3 Variation in shared encoding

Parallel to split encoding, shared encoding of nominal and locational predication can be attested in three variants. Of these variants, the "lexical" form, which involves the use of the same lexical item for nominal copula and locational support verb, is by far the most frequent. In addition to English, some other examples of this variant are Miskito (Misumalpan; Nicaragua) and Luganda (Bantu; Uganda).

(7) Miskito (Anonymous 1985: 213; Conzemius 1929: 110)

a. *Giovanni tuktan sirpi kum sa*
Giovanni child small one COP.3SG.PRES
'Giovanni is a small child.'

b. *aisi-kam bāra sa*
father-your here be.3SG.PRES
'Your father is here.'

(8) Luganda (Ashton et al. 1954: 434, 82)

a. *Mukasa n-ange tu-li babazzi*
Mukasa and-1SG 1PL.PRES-COP carpenters
'Mukasa and I are carpenters.'

b. *omugaati* *gu-li* *mu* *kabada*
loaf 3SG.PRES-be in cupboard
'The loaf is in the cupboard.'

The other two possible forms of shared encoding are rather uncommon. This is due to the fact that, for locational predication, the use of a full locational support item is the overwhelmingly preferred option (see Stassen 1997: 55–61). Thus, we only rarely find that a language has share status on the basis of a zero-zero encoding. One such case is Pitjantjatjara (Pama-Nyungan; South Australia).

(9) Pitjantjatjara (Douglas 1959: 55, 81)

a. *wait* *ngalyayala*
man doctor
'The man is a doctor.'

b. *tjitji* *kutjara* *ngura-ka*
child two camp-at
'The two children are at camp.'

Finally, share status for a language is also possible on the basis of a verbal encoding for both nominal and locational predicates. Since verbal encoding is definitely a minor typological option for both of these predicate types, it follows that a verbal-verbal shared encoding will be very uncommon as well. This variant can be illustrated by Korku (Munda; central India).

(10) Korku (Drake 1903: 149, 132, 80)

a. *ing* *shene-ba*
1SG go-NPST
'I go/will go.'

b. *di* *dhega* *kad* *ojha-ba*
that stone heavy load-NPST
'That stone is a heavy load.'

c. *di* *ura-gen-ba*
it house-at-NPST
'It is at home.'

As was the case with split encoding, the variation among languages with share status has been ignored for the purposes of this map.

4 Mixed encoding

In the above discussion, the split–share distinction has been defined as a binary parameter, in a yes/no fashion, and it will be represented as such on the map. It must be noted, however, that this binary definition is a simplification in some respects. For one thing, many languages have not just one encoding item for nominal predicates and locational predicates; commonly, copulas and locational support items come in sets, and these sets usually **coincide only partially**, if they coincide at all. This situation can be illustrated by Dutch. This language has a set of copular items (such as *zijn* 'to be', *worden* 'to become', *lijken* 'to appear'), as well as a set of locational verbs (such as *zijn* 'to be', *liggen* 'to lie', *hangen* 'to hang', *staan* 'to stand', and *zitten* 'to sit'). Now, the only items creating an overlap between these two sets are *zijn* 'to be' and *blijven* 'to stay', which can be used for both nominal and locational predication; all the other items are specialized for one or the other of the two predicational functions. Furthermore, since the use of *zijn* in locational function is much more limited in Dutch than is the use of *be* in that function in English, one may well ask whether Dutch should not be considered as a split-language rather than as a share-language.

A second factor which tends to blur the distinction between split-languages and share-languages is the phenomenon of **copularization** of the locational support verb. In some languages, the locational support verb has (or has attained) a limited possibility to act as the copula in nominal predication, in addition to the "real" copula that the language has. This leads to the possibility of a double encoding for nominal predications. Examples are from Spanish and from Tamil (Dravidian; southern India).

(11) Spanish (Max Kerkhof, p. c.)

a. *Julia* *es* *enfermera*
Julia COP.3SG.PRES nurse
'Julia is a nurse.'

b. *Julia* *está* *de* *enfermera* *(en Madrid)*
Julia be.3SG.PRES PREP nurse (in Madrid)
'Julia is a nurse (in Madrid).'

(12) Tamil (Asher 1982: 49, 50, 51)

a. *avaru* *(oru)* *ɖakʈar*
he (one) doctor
'He is a doctor.'

b. *ippo* *oru* *ɖakʈar-aa* *taan* *irukkaraaru*
now one doctor-ADV EMPH be.3SG.HON.PRES
'Now he is a doctor.'

c. *Raaman* *tooʈʈatt-ille* *irukkaraan*
Raaman garden-in be.3SG.M.PRES
'Raaman is in the garden.'

In the large majority of relevant cases, this double encoding of nominal predications is connected with a clear semantic difference, which can be described in terms of the notion of Time Stability (see Givón 1984) or Permanency (see Stassen 1997). For example, in the Spanish examples given above, it must be understood that the (a) sentence (which has a form of the "real" copula *ser*) indicates permanent class membership, whereas the (b) sentence (which contains the "copularized" locational verb *estar*) must be interpreted as stating that the class membership is only temporary. In this latter case, the sentence might well be translated as "Julia works/acts as a nurse in Madrid".

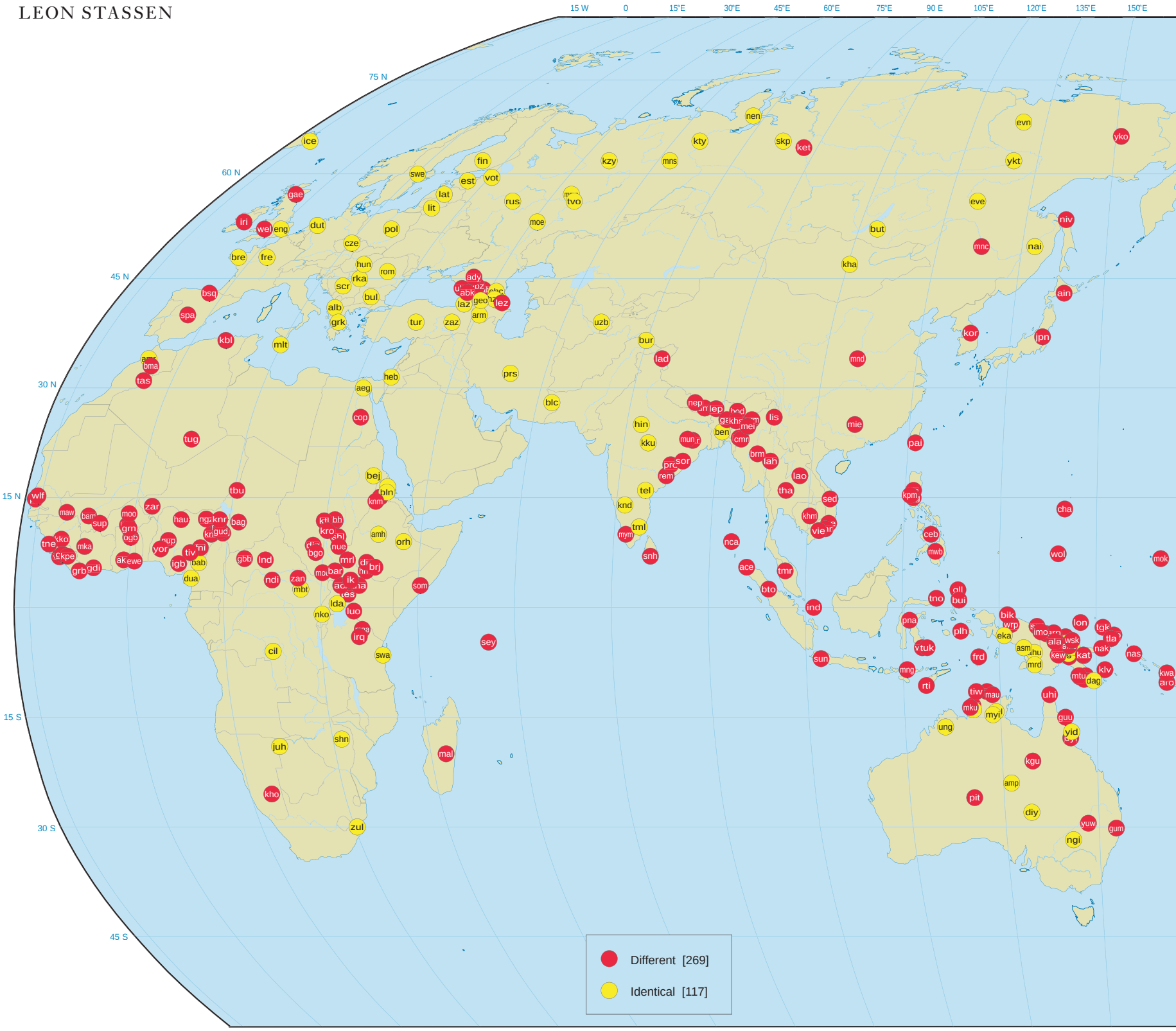
In view of the possible indeterminacy created by the phenomena of partial overlap and copularization, the map has been constructed along the following guidelines. A language will be called a share-language if at least one of its locational items can be used for copula function, unless this use is governed by conditions of Permanency. In all other circumstances, the language will be rated as a split-language. As a result, Dutch is rated as a share-language, while Spanish has been included among the split-languages.

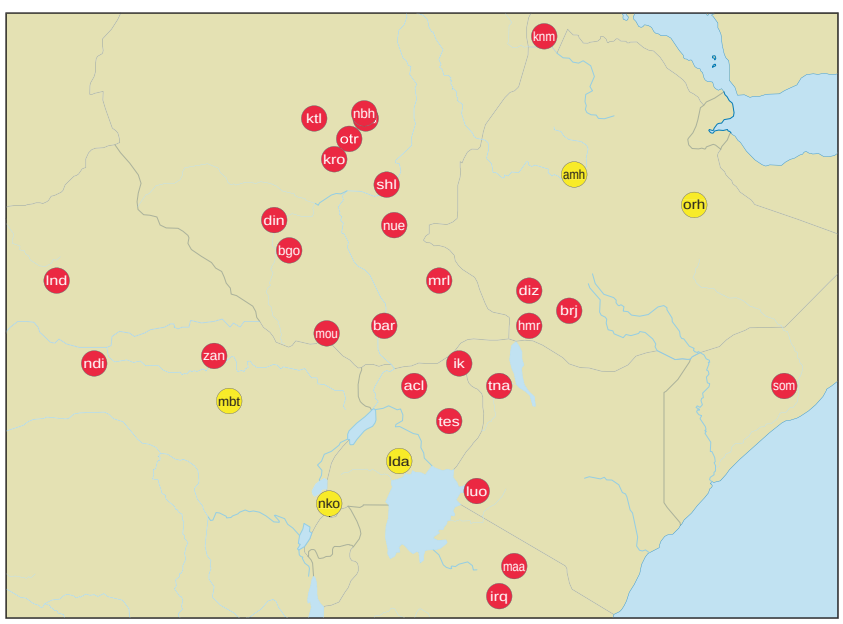
5 Geographical distribution

As the frequency numbers given above demonstrate, shared encoding is definitely the less frequent option among the world's languages. Nonetheless, there are a number of areas in which this encoding appears to be the rule. First, shared encoding is encountered in what might be called the Eurasian landmass, comprising Europe, central and northern Asia, the Middle East, Pakistan, and at least the northern part of India; notable exceptions here are some languages of the Caucasus and a number of languages on the western fringe of this macro-area (Celtic, Spanish). Secondly, shared encoding is prominent in Australia and New Guinea. Thirdly, we can note a concentration of shared encoding in an area which comprises the southern part of Central America and the northern part of South America. Finally, shared encoding is found in parts of eastern Africa, mainly due to the progress of copularization in the Bantu languages. Apart from these areas, however, split encoding seems to have a firm foothold.

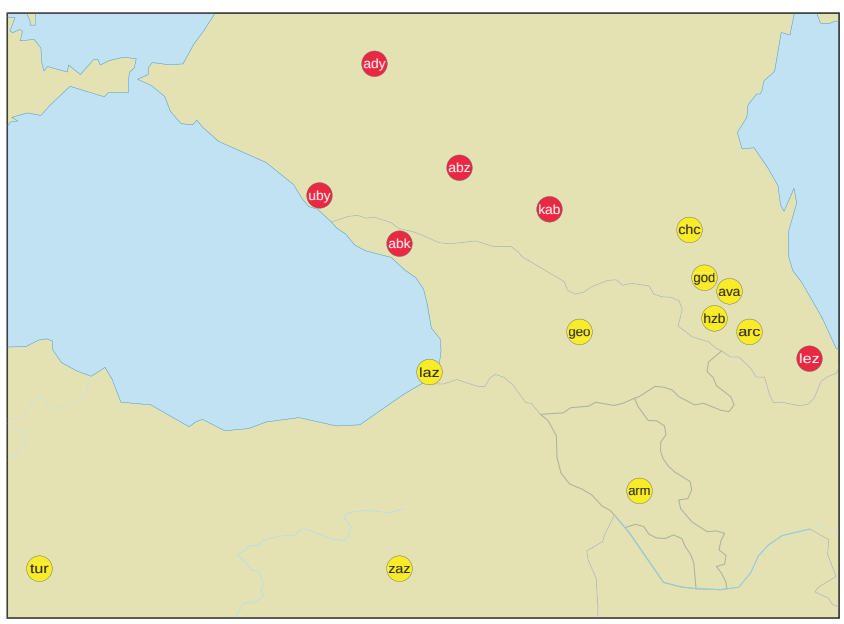
119 Nominal and Locational Predication

LEON STASSEN

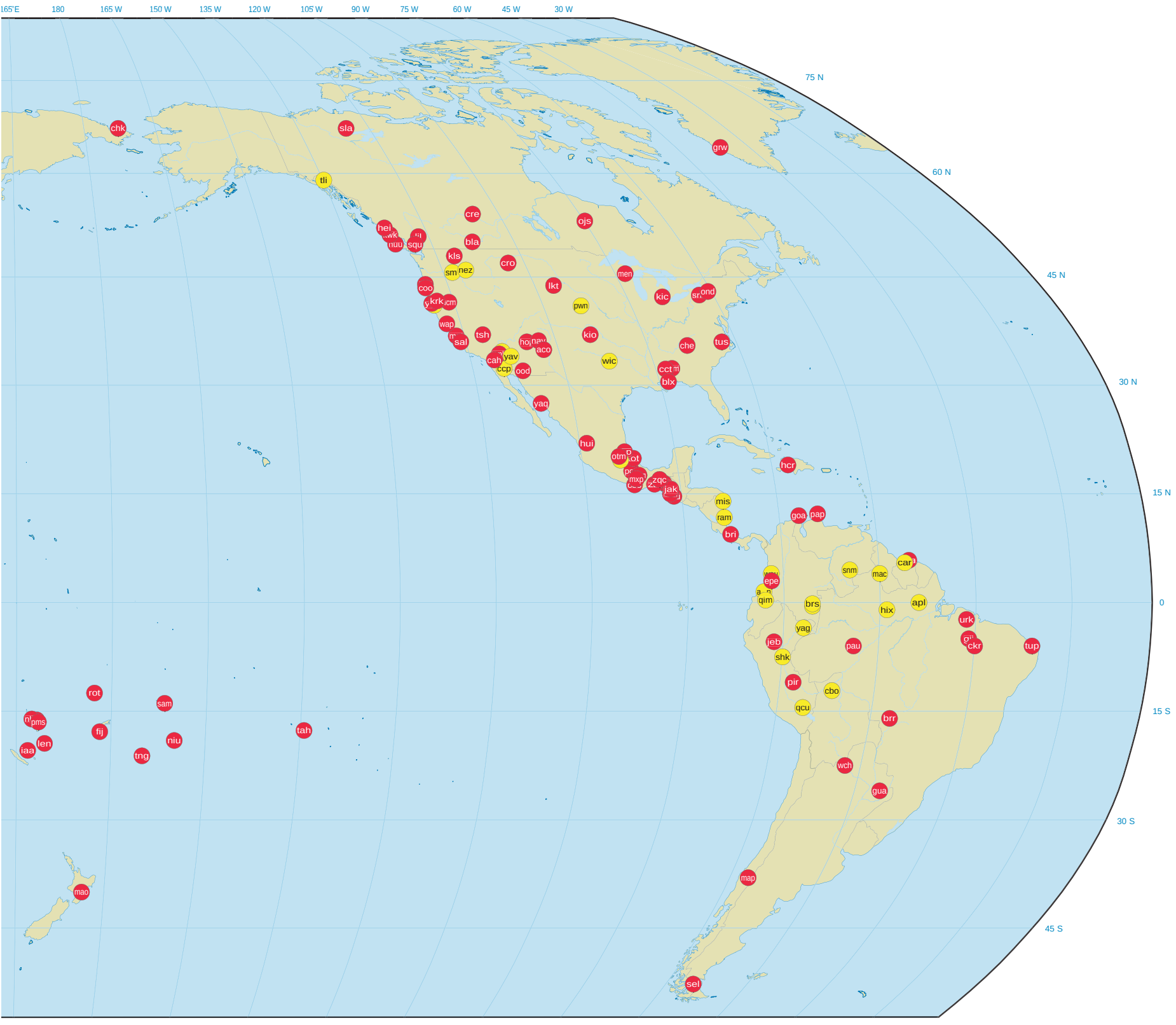




eastern Africa



Caucasus and adjacent areas



New Guinea and northern Australia

western United States

120 Zero Copula for Predicate Nominals

LEON STASSEN

1 Defining the values

This map shows the areal distribution of zero copula encoding for predicate nominals. That is, the map indicates whether a given language is like English, in which predicate nominals always require an overt copula (see 1), or rather like Russian, in which omission of the copula is allowed for at least some constructions (see 2).

(1) a. *John is a sailor.*
b. **John a sailor.*

(2) Russian (Maria Koptjevskaja-Tamm p. c.)
Moskva *gorod*
Moscow city
'Moscow is a city.'

Thus, the following values will be shown on the map:

1. Zero copula is impossible	211
2. Zero copula is possible	175
total	386

To avoid misunderstandings, a couple of remarks on the use of the term **zero copula** may be in order. This term has been used traditionally in the literature on predicate nominal sentences (see, for example, Benveniste 1966b). It will therefore be familiar to many readers, which is why I have decided to employ it in this chapter. I realize, however, that the term may give rise to Eurocentric interpretations, in that it might suggest that languages which do not have an overt copula are somehow "defective". On a more theoretical level, the term might suggest that, in languages with zero copula encoding of predicate nominal sentences, there is some phonologically null copula present in the syntax of these sentences, or, alternatively, that there is something corresponding to a copula in the conceptual/semantic structure underlying predicate nominal sentences in these languages. It should be stressed that none of these inferences are warranted. The term "zero copula" is used here as a strictly neutral technical label, in that it refers purely to a construction in which the relation between a subject and a nominal predicate is not marked by an overt item.

Furthermore, it should be noted that the scope of this chapter is restricted to stative predicate nominal sentences only. That is, the encoding of predicate nominal sentences which have a dynamic reading (commonly marked by some item meaning 'become') is not taken into consideration. Hence, when it is stated that a language has zero encoding in predicate nominal sentences, it should not be inferred that such languages cannot have an overt item in sentences which are similar to the English sentence *John became a dentist*; it only means that they lack an overt copular item in sentences that are similar to English *John is a dentist*.

2 The continuum of zero copula encoding

Although, for the purposes of this map, zero copula encoding is defined as a yes/no parameter, it is important to realize that this parameter in effect forms a continuum, with two extremes and a number of in-between cases. At one extreme, we have languages like English, in which zero copula encoding is never allowed. Other examples of this type are Finnish, Kabyle (Berber; Algeria), and Japanese.

(3) Finnish (Fromm and Sadeniemi 1956: 115)
ystävä-ni *on* *pappi*
friend-my COP.3SG.PRES vicar
'My friend is a vicar.'

(4) Kabyle (Hanoteau 1906: 88)
netsa *d'* *agellid'* *en* *temourth* *agi*
he COP king of country this
'He is king of this country.'

(5) Japanese (Makino 1968: 2)
John *wa* *sensei* *da.*
John TOP teacher COP
'John is a teacher.'

At the other end of the continuum we find languages in which zero copula encoding is mandatory for any construction that involves predicate nominals. Sinhala (Indic; Sri Lanka) and Tubu (Saharan; Libya) are cases in point. A concentration of languages of this type is found in Australia and New Guinea. Pitjantjatjara (Pama-Nyungan; Northern Territory, Australia) and Asmat (Trans-New Guinea; Papua, Indonesia) are examples.

(6) Sinhala (Gair 1970: 45)
unnæhee *huňgak* *prəsiddə* *kene-k*
he very famous person-NOM
'He is/was a very famous person.'

(7) Tubu (Lukas 1953: 170)
sígən *līfi*
3SG.EMPH orphan
'He is/was an orphan.'

(8) Pitjantjatjara (Douglas 1959: 55)
wait *ngalyayala*
man doctor
'The man is/was a doctor.'

(9) Asmat (Voorhoeve 1965a: 168)
no *ów* *akát*
1SG man handsome
'I am/was a handsome man.'

At intermediate positions on the continuum can be situated languages in which the use of the zero copula is limited by a condition of some sort. A well-known, and frequent, condition in this respect involves a split between present and nonpresent tense. Russian and Maltese are examples of languages in which a zero copula is used in the Present, whereas a full copula is mandatory for all other tenses.

(10) Russian (Maria Koptjevskaja-Tamm p. c.)
a. *ona* *vrač*
she doctor
'She is a doctor.'
b. *on* *byl* *učenik-om*
he be.M.PST pupil-INSTR
'He was a pupil.'

(11) Maltese (Albert Borg, p. c.)
a. *Albert* *tabib*
Albert doctor
'Albert is a doctor.'
b. *Albert* *kien* *tabib*
Albert be.3SG.M.PST doctor
'Albert was a doctor.'

An even more radically limited use of the zero copula is illustrated by Hungarian. Zero copula encoding in this language is restricted not only to the present tense, but also to sentences with third-person subjects. In all other cases, use of the full copula *vagy* is obligatory.

(12) Hungarian (Ginter and Tarnói 1991: 76, 78)

a. *én* *tanár* *vagyok*
1SG teacher be.1SG.PRES
'I am a teacher.'

b. *ö* *diák*
3SG.M pupil
'He is a pupil.'

For the purposes of the map, only languages of the English type have been rated as type one. In other words, a language is considered to be a member of type two if it allows the use of a zero copula, however minimal—as in the case of Hungarian—this use may be.

3 Types of copular items

As shown extensively by Stassen (1997: 76–100), not all items that are used as copulas in nominal predication have the same morphosyntactic status. A well-known class of copular items is formed by the **verbal copulas**, which have (by and large) the same morphosyntactic properties as verbs: English *be*, Spanish *ser*, and Russian *byt'* are illustrations of this class. In addition to this, however, large areas of the world show the use of nonverbal copular items. Particularly prominent is the use of a **pro-copula**, i.e. a demonstrative or personal pronoun which serves as the linker between subject and predicate nominal, and which is obligatory in nominal predication. We find this option especially in northern and central Asia, in North Africa and the Middle East, and in eastern Indonesia and Melanesia. Hebrew, Motu (Oceanic; Papua New Guinea), Turkish, and Beja (Cushitic; Sudan) are examples of languages with the pro-copula strategy; in Turkish and Beja, the pro-copula takes the form of a suffix.

(13) Modern Hebrew (Li and Thompson 1977: 428)
Moše *hu* *student*
Moshe 3SG.M student
'Moshe is a student.'

(14) Motu (Lister-Turner and Clark 1930: 54)
lau *na* *tau* *ia* *be* *hahine*
I DEM/COP man she DEM/COP woman
'I am a man, she is a woman.'

(15) Turkish (Lewis 1967: 98)
Türk-üm
Turk-1SG
'I am a Turk.'

(16) Beja (Tucker and Bryan 1966: 543)
barú:k *hadʔá:-bwa*
2SG.M sheik-2SG.M
'You are a sheik.'

In addition to pronouns, quite a few other nonverbal items are attested in copular function. Such **particle copulas** have their origin in a variety of markers of discourse-oriented phenomena such as topicalization, backgrounding, or contrastive focus for subjects or predicates. An example of such a particle copula was given in sentence (4), where the Kabyle copula *d'* can be identified as a general conjunctional/contrastive marker; for one thing, it is used as the conjunction 'and/with' between noun phrases. Another example is from Awtuw (Lower Sepik-Ramu; Papua New Guinea), where the item *po* is a general focus marker. Unlike in verbal sentences, the item is obligatory with nominal predicates. A similar situation holds in Vai (Mande; Liberia).

(17) Awtuw (Feldman 1986: 60, 138, 148)

a. *Yen* *w-æy-re*
2SG NONFACT-go-FUT
'You will go.'

b. *wan* *po* *ˈw-æy-rere*
1SG FOC NONFACT-go-DESID
'I am the one who wants to go.'

c. *wan* *po* *rameyæn*
1SG FOC/COP human.being
'I am a human being.'

(18) Vai (Welmers 1976: 116; Koelle 1854: 92)

a. *mànjáǎ* *mù* *ná'à* *à* *fɛ̀'ɛ̀*
chief FOC I him see.PERF
'It was the chief whom I saw.'

b. *mō* *kóro* *mú* *nda*
man old FOC/COP I
'I am an old man.'

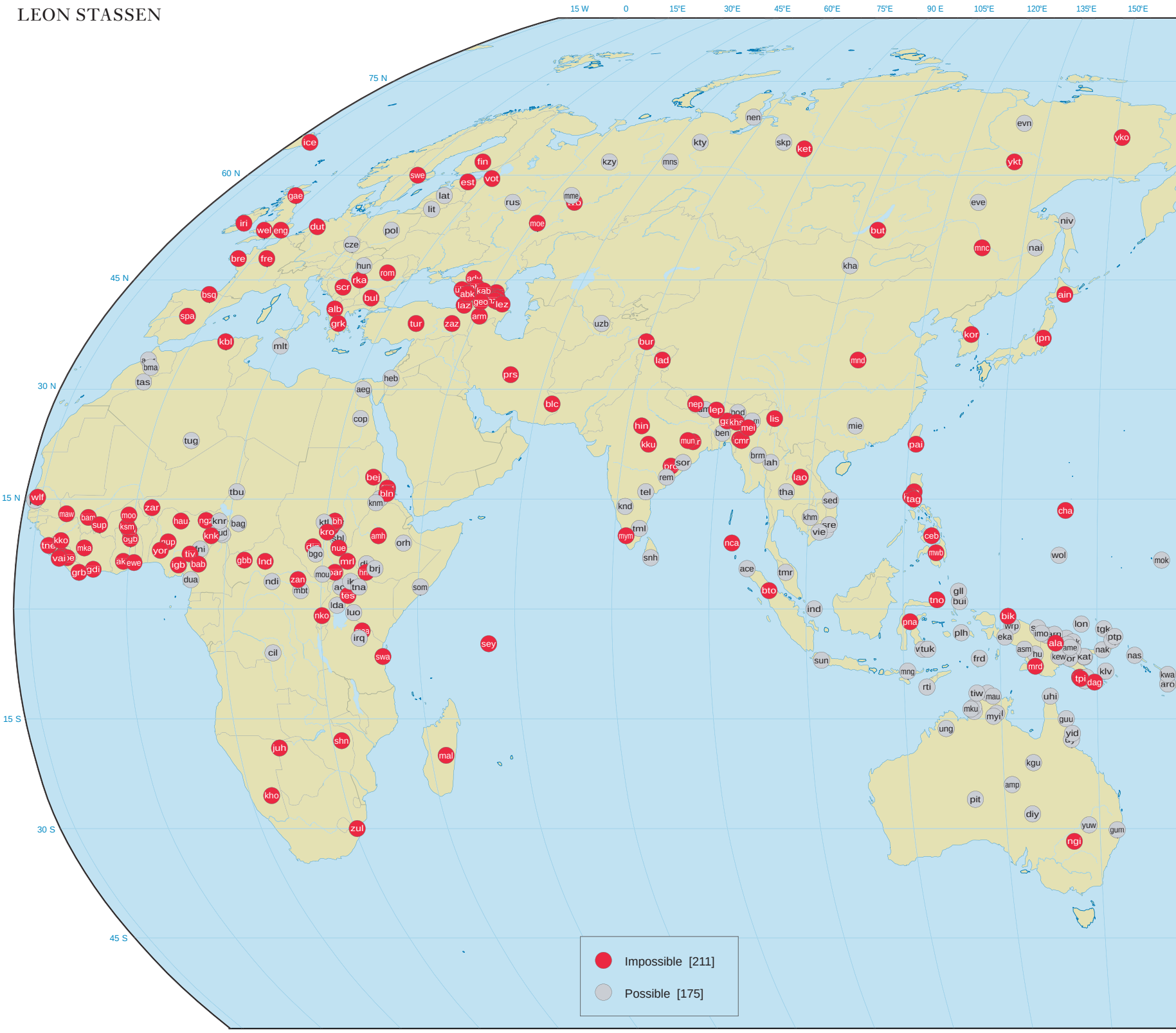
It must be concluded, then, that full copulas do not form a homogeneous morphosyntactic class. For the purposes of this map, however, differentiation between various classes of full copulas has been ignored. As a result, any language that has a full copula has been rated as a member of type one, regardless of the morphosyntactic status of the copular item (provided, of course, that the zero option is inadmissible in the language).

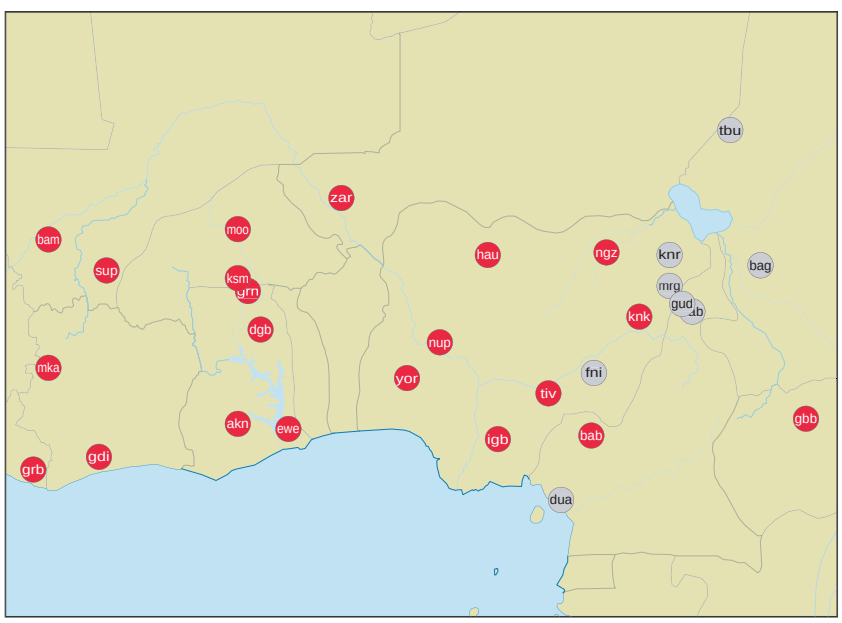
4 Geographical distribution

Regarding the distribution of zero copula encoding across the world's languages, the first thing to be noted is that there appears to be no major linguistic area in which the use of a zero copula is completely excluded. However, the areal distribution of zero copula encoding is far from random. The map shows a number of linguistic areas in which zero copula encoding is highly conspicuous. In particular, a zero strategy is used almost uniformly in the Pacific region (Australian, Papuan, and eastern Austronesian languages), and it is highly prominent in South America, as well as in the northern part of Africa. Opposed to this, we find areas in which a zero copula constitutes a minor or marginal option. This is the case in Europe and the Middle East (covering Indo-European, Semitic, Turkic, Caucasian, and Uralic languages), northern India and the Himalayas, sub-Saharan Africa, North and Central America, and East Asia (covering China, Korea, Japan, and the Philippines). It should be added that, quite often, the distribution of zero copulas within a language family or area is rather unpredictable. Languages which are closely related areally or genealogically may differ considerably in the extent to which they allow zero encoding. For instance, while Austronesian languages typically opt for zero encoding, a full copula is mandatory in a closely related group of three Austronesian languages from northern and central Vanuatu (Lonwolwol, Big Nambas, and Paamese).

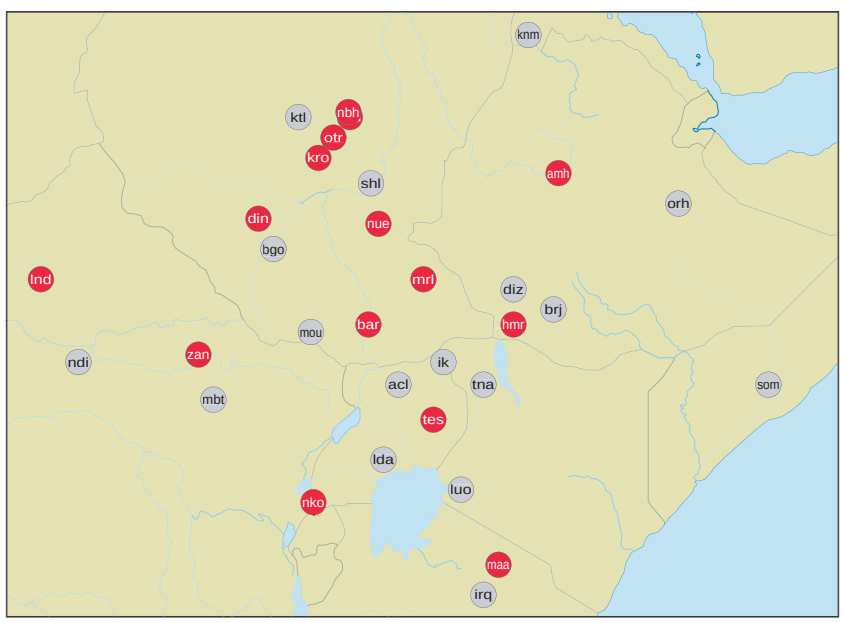
120 Zero Copula for Predicate Nominals

LEON STASSEN

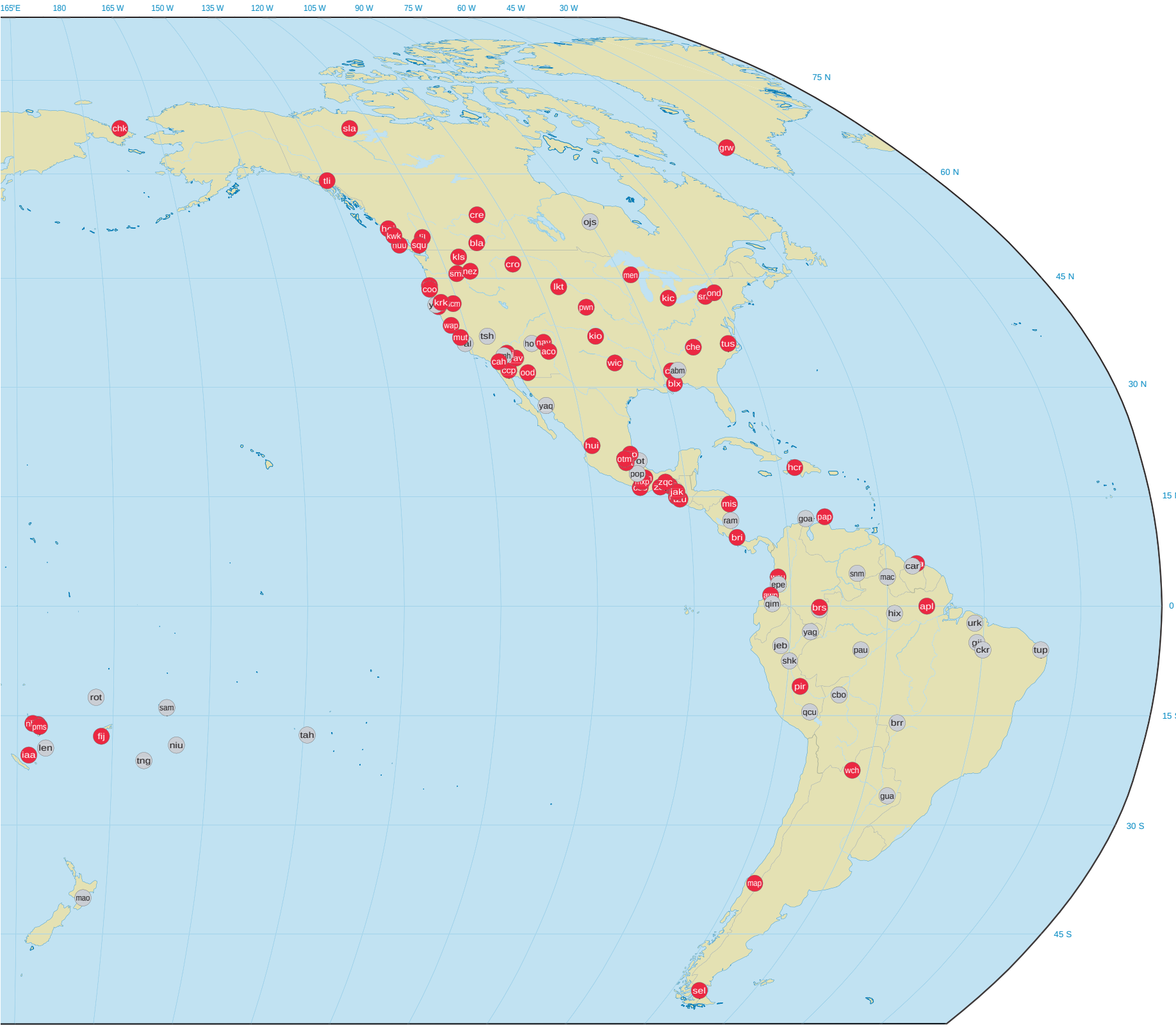




Nigeria and surrounding areas



eastern Africa



New Guinea and northern Australia

western United States

121 Comparative Constructions

LEON STASSEN

1 Defining the construction

In semantic or cognitive terms, comparison can be defined as a mental act by which two objects are assigned a position on a predicative scale. If the positions on the scale are different, then we speak of **the comparison of inequality**, which finds its linguistic encoding in **comparative constructions**. Such a construction essentially involves three things: a predicative scale, which, in language, is usually encoded as a gradable predicate, and two objects. Although these objects can, in principle, be complex, the practice of typological linguistic research has been to restrict them to primary objects, which are typically encoded in the form of noun phrases. Thus, a comparative construction typically contains a predicate and two noun phrases, one of which is the object of comparison (**the comparee NP**), while the other functions as the "yardstick" of the comparison (**the standard NP**). In short, prototypical instances of comparative constructions in the languages of the world are sentences that are similar to the English sentence in (1), in which the noun phrase following the item *than* is the standard NP:

(1) *John is taller than Lucy.*

2 Types of comparative constructions

Modern literature on the typology of the comparison of inequality includes Ultan 1972, Andersen 1983, and Stassen 1984, 1985. The last of these authors presents a typology of comparative constructions which boils down to four major types. A basic parameter in this typology is the encoding of the standard NP. First, one can make a distinction between instances of **fixed-case comparatives** and **derived-case comparatives**. In the former type, the standard NP is always in the same case, regardless of the case of the comparee NP. In the latter type, the standard NP derives its case assignment from the case of the comparee NP. Classical Latin is an example of a language in which both types are allowed. The sentences in (2) illustrate a construction type in which the standard NP is dependent on the comparee NP for its case marking: it can either be in the nominative case (*tu*) or in the accusative case (*te*). In contrast, sentence (3) shows a construction type in which the standard NP is invariably in the ablative case (*te*). As a result, sentence (3) is ambiguous between the readings of (2a) and (2b).

(2) Classical Latin (Kühner and Stegmann 1955: 466)

a. *Brutum* *ego* *non* *minus* *amo*
Brutus.ACC 1SG.NOM not less love.1SG.PRES
quam *tu*
than 2SG.NOM
'I love Brutus no less than you (love Brutus).'

b. *Brutum* *ego* *non* *minus* *amo*
Brutus.ACC 1SG.NOM not less love.1SG.PRES
quam *te*
than 2SG.ACC
'I love Brutus no less than (I love) you.'

(3) Classical Latin (Kühner and Stegmann 1955: 466)
Brutum *ego* *non* *minus* *te* *amo*
Brutus.ACC 1SG.NOM not less 2SG.ABL love.1SG.PRES

Both types of comparative constructions can be subcategorized further, on the basis of additional parameters. Within the fixed-case comparatives, a first distinction is that between **Exceed Comparatives** and **Locational Comparatives**. **Exceed Comparatives** have as their characteristic that the standard NP is constructed as the direct object of a transitive verb with the meaning 'to exceed' or 'to surpass'. Thus, the construction typically includes two predicates, one which is the comparative predicate, and another which is the 'exceed'-verb. The comparee NP is the subject of the 'exceed'-verb. Duala (Bantu; Cameroon) presents an instance of the Exceed Comparative, as does Thai.

(4) Duala (Ittmann 1939: 187)
níǹ *ndábò* *e* *koló* *búkà* *nínɛ̰*
this house it big exceed that
'This house is bigger than that.'

(5) Thai (Warotamasikkhadit 1972: 71)
kăw *sŭuŋ* *kwă* *kon* *túk* *kon*
he tall exceed man each man
'He is taller than anyone.'

Locational Comparatives, on the other hand, are characterized by the fact that the standard NP is invariably constructed in a case form which also has a locational/adverbial function. Depending on the exact nature of this function, Locational Comparatives can be divided into three further subtypes. *From*-comparatives mark the standard NP as the source of a movement, with a marker meaning 'from' or 'out of'. *To*-comparatives construct the standard NP as the goal of a movement ('to, towards', 'over, beyond') or as a benefactive ('for'). Finally, *At*-comparatives encode the standard NP as a location, in which an object is at rest ('in', 'on', 'at', 'upon'). Illustrations of the various subtypes of Locational Comparatives are from Mundari (Munda; India), Uzbek (Turkic; Uzbekistan), and Estonian for *from*-comparatives; Siuslaw (Siuslawan; Oregon) and Maasai (Nilotic; Kenya and Tanzania) for *to*-comparatives; and Ahaggar Tuareg (Berber; southern Algeria) and Tubu (Nilo-Saharan; Chad and Niger) for *at*-comparatives. For the purposes of the map, however, this internal variation within the Locational Comparative has been ignored.

(6) Mundari (Hoffmann 1903: 110)
sadmo-ete *hati* *maranga-e*
horse-from elephant big-3SG.PRES
'The elephant is bigger than the horse.'

(7) Uzbek (Sjoberg 1963: 142)
ɔtam *u* *ɔdam-dan* *yɔš*
father.my that man-from young
'My father is younger than that man.'

(8) Estonian (Oinas 1966: 140)
kevad *on* *sügis-est* *ilusam*
spring is fall-from more.beautiful
'The spring is more beautiful than the fall.'

(9) Siuslaw (Frachtenberg 1922a: 555)
sea *his* *na-tc*
he good me-to
'He is better than me'

(10) Maasai (Tucker and Mpaayi 1955: 93)
sapuk *olkondi* *to* *lkibulekeny*
big hartebeest to waterbuck
'The hartebeest is bigger than the waterbuck.'

(11) Ahaggar Tuareg (Hanoteau 1896: 52)
kemmou *tehousid* *foull* *oult* *ma* *m*
you pretty.2SG.F upon sister of you
'You are prettier than your sister.'

(12) Tubu (Lukas 1953: 45)
sa-umma *gere* *do* *mado*
eye-his blood on red
'His eye is redder than blood.'

Turning now to the derived-case comparatives, in which the case marking of the standard NP is derived from—or "parasitic on"—

the case marking of the comparee NP, we note that, once again, two subtypes can be distinguished. First, there is the **Conjoined Comparative**. Here the comparative construction usually consists of two structurally independent clauses, one of which contains the comparee NP, while the other contains the standard NP. Furthermore, the two clauses show a structural parallelism, in that the grammatical function of the comparee NP in one of the clauses is duplicated by the grammatical function of the standard NP in the other clause. If, for example, the comparee functions as the grammatical subject in its clause, the standard NP will also have subject status in its clause.

Since the construction has two clauses, it follows that the construction will also have two independent predicates. In other words, the comparative predicate is expressed twice. There are two ways in which this double expression may be effectuated. The language may employ antonymous predicates in the two clauses ('good-bad', 'strong-weak'). Alternatively, the two predicates may show a positive-negative polarity ('good-not good', 'strong-not strong'). An example of the first variant is found in Amele (Madang; north-eastern Papua New Guinea); the second variant has been attested for Menomini (Algonquian; Wisconsin). Sentence (15) illustrates one of the comparative constructions in Malay. Here the standard-NP and the comparee-NP are conjoined as sentence topics, and the following clause predicates the property of the comparee-NP only; that is, in this (rather infrequent) variant of the Conjoined Comparative the comparative predicate is expressed only once. For the purposes of the map, all variants of the Conjoined Comparative are treated as a single category.

(13) Amele (Roberts 1987: 135)
jo i ben jo eu nag
house this big house that small
'This house is bigger than that house.'

(14) Menomini (Bloomfield 1962: 506)
Tata'hkes-ew nenah tɛh kan
strong-3SG I and not
'He is stronger than me.'
(lit. 'He is strong and I (am) not (strong).')

(15) Malay (Lewis 1968: 157)
kayu batu bĕrat batu
wood stone heavy stone
'Stone is heavier than wood.'

A second subtype of derived-case comparison is defined negatively, in that the standard NP has derived case, but the construction does not have the form of a coordination of clauses. Instead, the construction features a specific **comparative particle** which accompanies the standard NP. The English *than*-comparative is an instance of this **Particle Comparative.** Other examples are the comparative construction in French, with its comparative particle *que*, and the comparative construction in Hungarian, which features the particle *mint* 'than, like'.

(16) French (Bernard Bichakjian, p. c.)
tu es plus jolie que ta sœur
you are more pretty than your sister
'You are prettier than your sister.'

(17) Hungarian (Edith Moravcsik, p. c.)
István magasa-bb mint Peter
István NOM tall-more than Peter.NOM
'István is taller than Peter.'

In summary, the map of comparative constructions shows the areal distribution of four types. Two of these types (viz. the Locational Comparative and the Exceed Comparative) are instances of fixed-case comparison, while the other two (viz. the Conjoined Comparative and the Particle Comparative) are instances of derived-case comparison. As a result, the following four values are shown on the map:

1.	Locational Comparative	78
2.	Exceed Comparative	33
3.	Conjoined Comparative	34
4.	Particle Comparative	22
	total	167

3 Predicate marking in comparative constructions

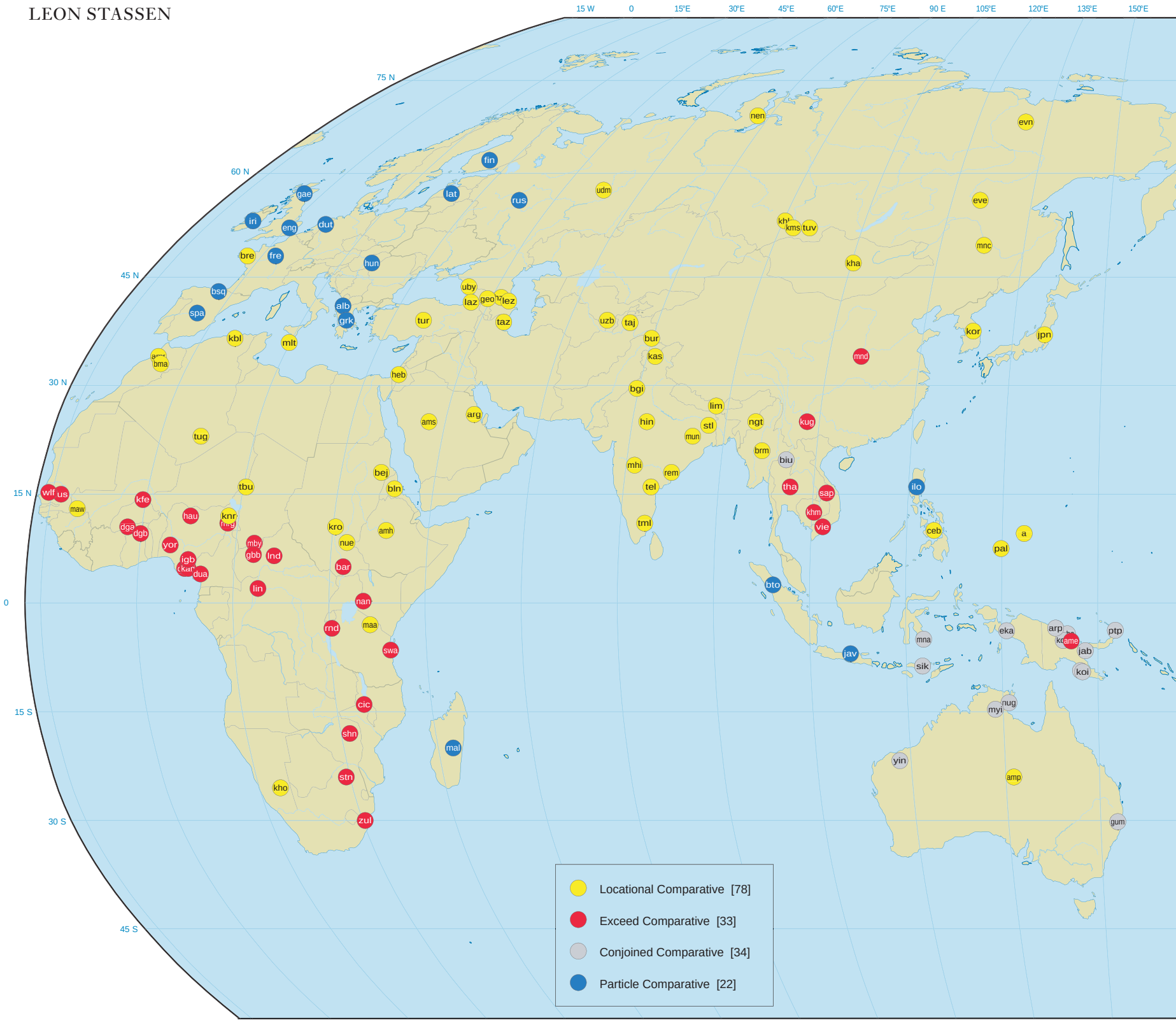
Apart from, or in addition to, case assignment of the standard NP, a further possible parameter in the typology of comparative constructions might be considered to be the presence or absence of comparative marking on the predicate. In the vast majority of languages, such overt marking is absent; predicative adjectives in comparatives retain their unmarked, "positive" form. Some languages, however, mark a predicative adjective in a comparative construction by means of a special affix (e.g. *-er* in English, German, and Dutch, *-ior* in Latin, *-bb* in Hungarian, *-ago* in Basque) or a special adverb (*more* in English, *plus* in French). Especially in the case of comparative affixes, the etymological origin is largely unknown. As for the areal distribution of predicate marking in comparatives, it is an almost exclusively European phenomenon, and is particularly frequent in languages that have a particle comparative construction. For a tentative explanation of this latter correlation see Stassen (1985: ch. 15). In the map the phenomenon of comparative predicate marking has not been taken into account.

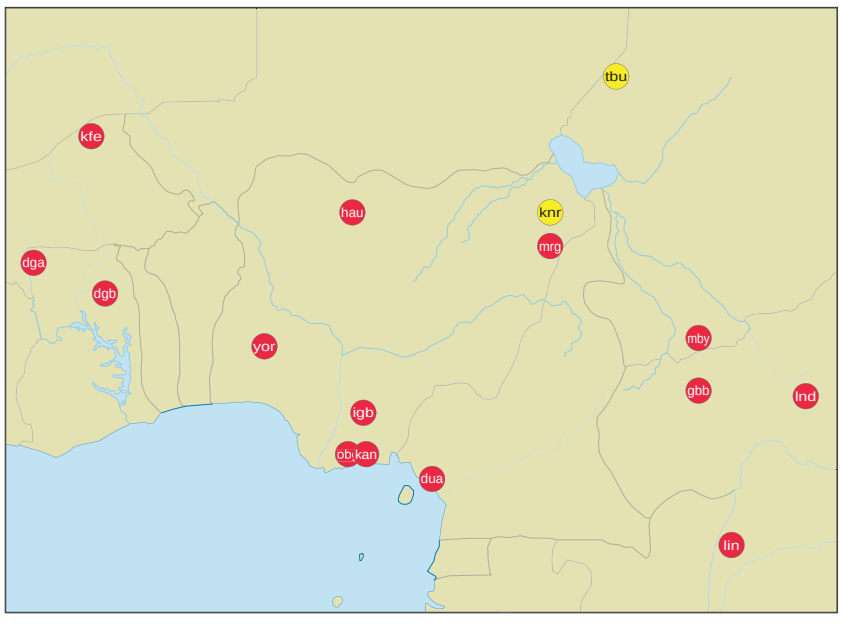
4 Geographical distribution

As the map demonstrates, the areal distribution of the various types of comparative constructions is striking (see also Heine 1994). For one thing, the Exceed Comparative appears to be almost exclusively restricted to two areas, viz. sub-Saharan Africa, and China and South-East Asia. No less limited is the distribution of the Particle Comparative, which turns out to have its base in the modern languages of Europe; instances of this type outside Europe (such as the Uto-Aztecan languages in North and Central America) may well be cases of influence from English and/or Spanish. The Conjoined Comparative has a stronghold in Australia and New Guinea, and is also prominent in the Amazon basin. Finally, the Locational Comparative is the rule in northern Africa and in the vast landmass of Eurasia (including the Middle East and India, but excluding Europe), and can also be found in Eskimo languages and in scattered instances over the Americas, in Polynesia, and in Australia and Papua New Guinea.

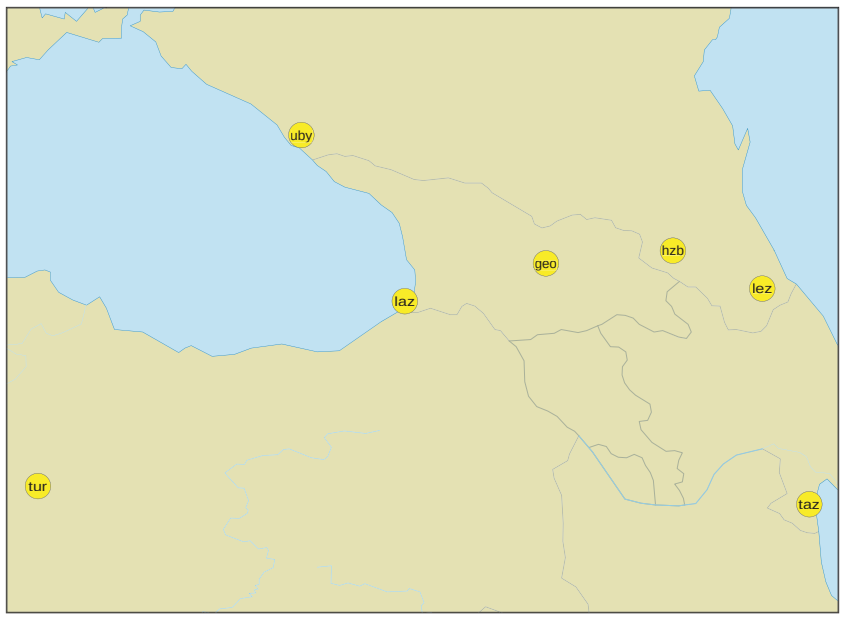
121 Comparative Constructions

LEON STASSEN

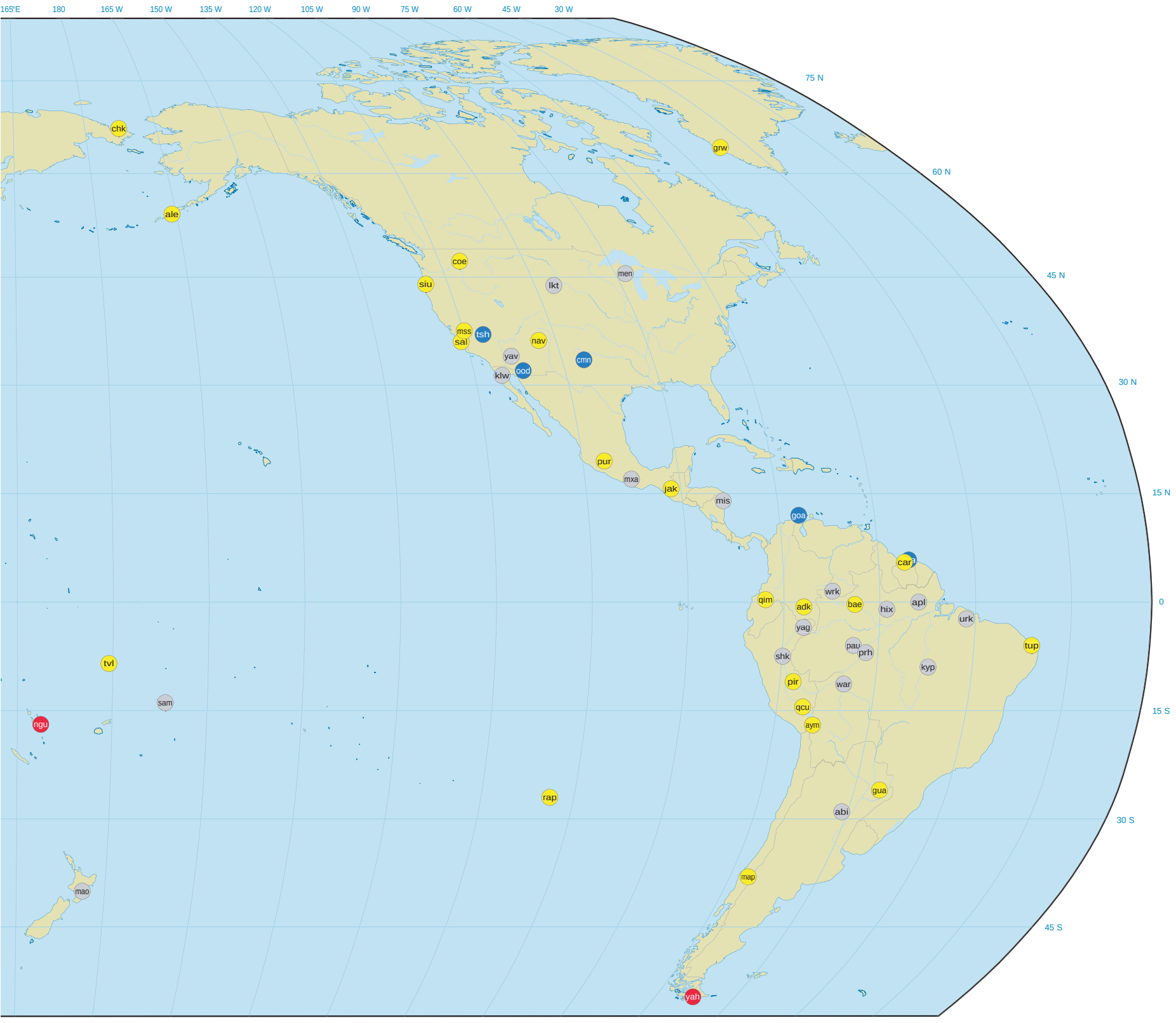


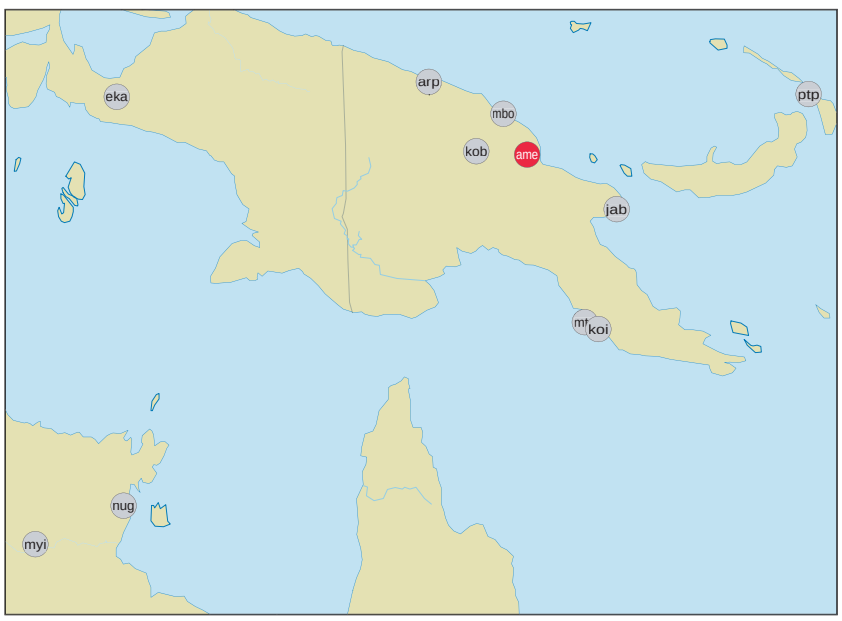


Nigeria and surrounding areas

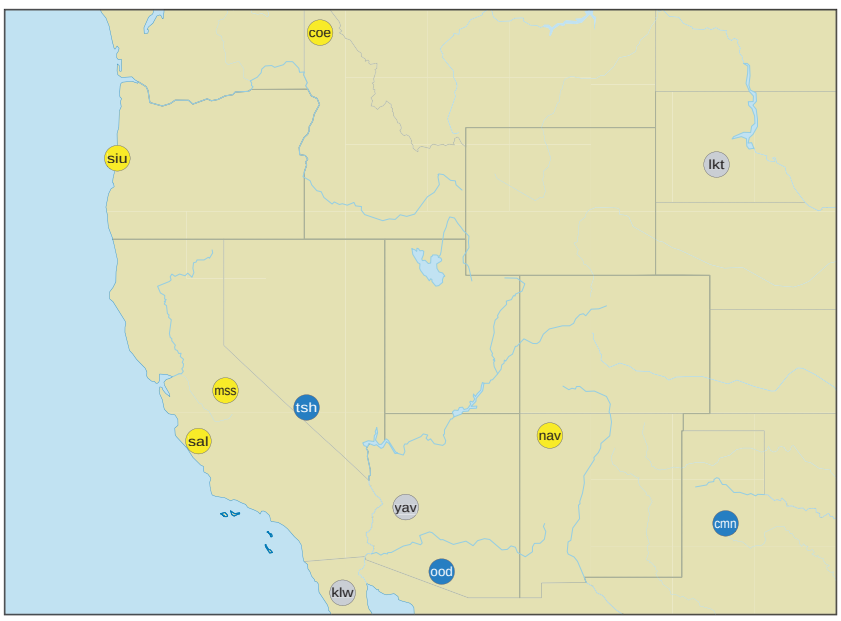


Caucasus and adjacent areas





New Guinea and northern Australia



western United States

122–123 Relativization Strategies

BERNARD COMRIE AND TANIA KUTEVA

1 Introduction

A **relative clause** is a clause narrowing the potential reference of a referring expression by restricting the reference to those referents of which a particular proposition is true. Thus, along with the main clause *I teach the girl . . .* , the English sentence in (1) contains a relative clause, *who just greeted us*, which narrows the potential reference of the referring expression, *the girl*—called the head noun—to only referents of which the proposition *(the girl) just greeted us* is true.

(1) *I teach the girl who just greeted us.*

The relative clause and its head noun form the **relative construction**. For the present maps, we have not taken into account relative clauses which have no head noun (like English *what you don't know*). Languages use different strategies to encode the relative construction; we will refer to these as **relativizing strategies**.

There are different perspectives from which relativizing strategies can be studied. Thus, from the point of view of the **linear order** of the head noun and the relative clause, we can distinguish prenominal, postnominal, and circumnominal embedded relative clauses as well as preposed, postposed, and adjoined relative clauses (Lehmann 1984; and see Chapter 90). Another possible perspective involves the **global cognitive mechanisms** underlying relativizing strategies on a language-universal level; from this perspective we can distinguish two global strategies, combining and inserting (Kibrik 1992).

For the purposes of the present study, we classify our sample languages according to the mechanisms by which the language in question expresses the syntactic-semantic role of the head noun in the relative clause, whereby we consider only formally expressed morphosyntactic means. In (1), for instance, the head noun serves as subject of the relative clause, and this is marked in English by use of the nominative relative pronoun *who*.

Languages may employ different morphosyntactic (as well as suprasegmental) means, that is, different relativizing strategies, for different syntactic–semantic roles of the head noun. In the English sentence in (1), the head noun has the subject role, and it is relativized by means of a relative pronoun. If the same head noun, *the girl*, has the role of the object, one of the ways in which it may be relativized is by not using any morphosyntactic (and/or suprasegmental) element at all, i.e. by means of a "gap" (see §2 below), as in (2):

(2) *the girl we saw yesterday*

This latter possibility does not exist when the role of the head is subject; hence (3) is ungrammatical in (standard) English.

(3) **The girl [just greeted us] is a student of mine.*

Since many languages use different strategies for relativizing on different roles, we distinguish between **relativization on subjects** and **relativization on obliques**. Map 122 shows what strategies the languages of the world use to relativize on the subject.

For Map 123, we take into consideration the relativizing strategies the languages of the world employ with obliques, whereby we take the instrumental to be the prototypical case of obliques. In languages where we have no ready access to information about relativization on instrumentals, we consider other, comparable constructions: the comitative, the indirect object, the benefactive, the locative, etc., but *not* the possessive or the temporal.

Note that the head noun may have different roles in the main clause and in the relative clause. Thus in the English sentence in (4), the head noun *the girl* functions as the object of the main clause, and the subject of the relative clause.

(4) *I like the girl who greeted us yesterday.*

For the purposes of both Map 122 and Map 123, the role relevant to our classification is the one the head noun has **within** the relative clause. Accordingly, the example in (4) is an example of relativizing on the subject, since the head noun *the girl* functions as the subject of the relative clause.

The present classification is based on the assumption that *all* natural languages can relativize on subjects; hence, we are including multipurpose clauses like those in the Diyari (Pama-Nyungan; South Australia) example (5) which have several other functions apart from the function of relativization. Diyari has no specific subordination construction whose sole, or even prototypical, function is to encode a relative clause. It uses, instead, a general, unified modifying construction which—depending on context—may be interpreted as either a subordinate temporal, conditional or relative clause, as in (5).

(5) Diyari (Austin 1981: 209)

ṭanali	*ṉiṉa*	*waḷa*	*ŋanka-ṇa*	*ṭaḷaṛa*
3PL.ERG	3SG M.ACC	nest-ABS	make-REL.SS	rain

maḍa	*kuda-ṇa*	*ŋaṛi-yi*	*waḷa-ṉi*
stone-ABS	put-PCL	go.down-PRES	nest-LOC

'If/when/after they make/made the nest, they put the rain stone in it.'
'Having made the nest, they put the rain stone in it.'
'They who make/made the nest put the rain stone in it.'
'They put the rain stone in the nest they make/made.'

Instead of regarding languages such as Diyari as irrelevant to relativizing strategies, in the present classification we treat them as languages using a gap in relativizing on the subject role.

2 Defining the values

2.1 Map 122 Relativization on subjects. Map 122 shows the distribution of these four relativizing strategies with subjects:

1. Relative pronoun	12
2. Nonreduction	24
3. Pronoun retention	5
4. Gap	125
total	166

The first strategy has come to be called the **relative pronoun strategy**: the position relativized is indicated inside the relative clause by means of a clause-initial pronominal element, and this pronominal element is case marked (by case or by an adposition) to indicate the role of the head noun within the relative clause. The German example in (6) illustrates the relative pronoun strategy with subjects.

(6) German

Der Mann,	*[der*	*mich*	*begrüßt*	*hat],*
man.NOM	REL.NOM	me	greet.PTCP	has

war	*ein*	*Deutscher.*
be.3SG.PST	one	German

'The man who greeted me was a German.'

Note that the mere presence of a pronoun that is restricted to relative clauses, and is thus in some intuitive sense a relative *pronoun*, is not sufficient to define an instance of the pronoun *strategy* (Comrie 1998: 61–2). Such a relative pronoun can be case marked, for instance, not to indicate its role in the relative clause, but rather to

agree in case with the head noun in the main clause. Thus in the Modern Standard Arabic sentence in (7), the relative pronoun is nominative, like the head noun, whereas the position relativized in the relative clause is direct object (which would require the accusative case in Arabic); such instances do not satisfy our definition of the relative pronoun strategy.

(7) Modern Standard Arabic (Comrie 1998: 62)
ʾal-ɣulaam-aani *l-musiiqiyy-aani* *llaðð-aani*
the-boy-DU NOM the-musical-DU.NOM REL-DU NOM
'the two boy musicians (whom Cyrano sent . . .)'

The second major relativizing strategy identifiable in our language sample is the **nonreduction strategy**: the head noun appears as a full-fledged noun phrase within the relative clause. Comrie (1989) and Comrie (1998) distinguish two subtypes of this strategy, which we have also been able to identify in our sample languages. The first subtype involves **correlative clauses**, where the head noun appears as a full-fledged noun phrase in the relative clause and is taken up again by a pronominal or a nonpronominal element in the main clause; this subtype is exemplified by the Pirahã (Mura; Brazil) sentence in (8).

(8) Pirahã (Everett 1986: 276)
boitóhoi *bog-ái-hiab-i-s-aoaxái*
boat come-ATELIC-NEG-EPENTH-ʔ-INTER
boitó *báosa* *xig-i-sai* *(híx)*
boat barge bring-EPENTH-NMLZ (COMP/INTER)
'Might it be that the boat (which) tows barges is not coming?'

The second subtype of the nonreduction strategy, **internally headed relative clauses**, covers cases where the head is represented by a full noun phrase inside the relative clause, and has no explicit representation in the main clause, as exemplified by the Maricopa (Yuman; Arizona) sentence in (9).

(9) Maricopa (Gordon 1986: 255)
aany=lyvii=m *ʼiipaa* *ny-kw-tshqam-sh* *shmaa-m*
yesterday man 1-REL-slap.DIST-SUBJ sleep-REAL
'The man who beat me yesterday is asleep.'

In addition to the above two subtypes of the nonreduction strategy distinguished in Comrie (1989; 1998), in the present classification the nonreduction strategy covers one more subtype that we have termed elsewhere (see Kuteva and Comrie, forthcoming) the **paratactic relative clause** (cf. English *That man just passed by us, he introduced me to the Chancellor of the University yesterday*). This subtype has the following characteristics: the "relative" clause contains the full-fledged head and is the same as an unmarked simple (declarative) clause; the relative and main clauses are only very loosely joined together. The Amele (Trans-New Guinea; Papua New Guinea) sentence in (10) illustrates this subtype of the non-reduction relativizing strategy.

(10) Amele (John Roberts, p. c.)
mel *mala* *heje* *on* *((mel)*
boy chicken illicit take.3SG.SUBJ-REM.PST boy
eu) *busali* *nu-i-a*
that run.away go-3SG.SUBJ-TOD.PST
'The boy that stole the chicken ran away.'

In (10), *mel* 'boy' is the "relativized" noun in the "relative" clause. This nominal can be optionally referred to in the following "main" clause either by the demonstrative *eu* 'that' or, if clarification is needed, by *mel eu* 'boy that'. What links the two clauses is the rising intonation at the end of the first clause. This indicates that it is not a final clause and is in either a subordinate or coordinate relationship with the following clause. The difference between correlatives and paratactic relatives is that whereas in correlatives, the relative clause usually contains some element (e.g. an interrogative) which would not be present in the corresponding simple declarative sentence, paratactic relatives contain no such element.

The third major relativizing strategy with subjects is the **pronoun-retention strategy**. In languages employing this strategy, the position relativized is explicitly indicated by means of a resumptive personal pronoun, as in the case of the Babungo (Bantu; Cameroon) example in (11).

(11) Babungo (Schaub 1985: 34)
mə̀ *yè* *wə́* *ntɨ́ə* *fáŋ* *ŋwə́* *sɨ́* *sàŋ* *ghə̂*
I see.PFV person that who he PST2 beat.PFV you
'I have seen the man who has beaten you.'

Note that we define the pronoun retention strategy as one where a pronoun or pronominal marker referring to the head of a relative clause is obligatory in the relative clause but is not obligatory in the corresponding simple declarative clause. Thus example (11) is considered pronoun-retention because it obligatorily includes the pronoun glossed 'he', although such a pronoun does not occur in the simple sentences in (12).

(12) Babungo (Schaub 1985: 23)
a. *Làmbí* *ˈsáŋ* *ŋwə́*
Lambi beat.IMPF him
'Lambi beat him.'
b. *Làmbí* *sáŋ* *ˈŋwə́*
Lambi beat.PFV him
'Lambi has beaten him.'

Note also that, in contrast to Babungo, there exist languages where a pronominal subject marker is obligatorily attached to the verb both in the relative clause and in the corresponding simple declarative clause, as in Maybrat (West Papuan; Papua, Indonesia), for instance:

(13) Maybrat (Philomena Dol, p. c.)
a. *fai* *m-ait* *awiah*
woman 3.UNMARKED-eat taro
'The woman eats taro.'
b. *fai* *ro* *m-ait* *awiah*
woman REL 3.UNMARKED-eat taro
'the woman who eats taro'

In our classification, such languages are not treated as cases of the pronoun-retention strategy.

Finally, the fourth major relativizing strategy with subjects identifiable across languages is the **gap strategy**. This strategy involves cases where there is no overt case-marked reference to the head noun within the relative clause:

(14) Turkish (Comrie 1998: 82)
[kitab-ɪ *al-an]* *öğrenci*
book-ACC buy-PTCP student
'the student who bought the book'

Note that for present purposes the gap strategy unites a number of possibilities that would need to be kept apart for other purposes. In some languages, the gapped clause construction may be only one manifestation of a single formal means for marking not only what translates English relative clauses but also a number of other clause types, e.g. the Fact-S construction (as in "The fact that he doesn't know me . . ."), etc., where there is no gap. More precisely, this construction can be regarded as a *general noun-modifying clause construction*, as the following examples from Karachay-Balkar (Turkic; northern Caucasus) demonstrate:

(15) Karachay-Balkar (Comrie 1998: 81)
a. *[kitab-ï* *al-ɣan]* *oquwčču*
book-ACC buy-PTCP student
'the student who bought the book'
b. *[oquwčču* *al-ɣan]* *kitap*
student buy-PTCP book
'the book that the student bought'
c. *[prezident* *kel-gän]* *hapar*
president come-PTCP news
'the news that the president has come'
d. *[et* *biš-gän]* *iyis*
meat cook-PTCP smell
'the smell of meat cooking'

By contrast, there are languages with a gap strategy where different constructions must be used depending on the position relativized, on whether a Fact-S construction is involved, etc., cf. (16a) versus (16b–c):

(16) Turkish (Comrie 1998: 82)

a. *[kitab-ı al-an] öğrenci*
book-ACC buy-PTCP student
'the student who bought the book'

b. *[öğrenci-nin al-dığ-ı] kitap*
student-GEN buy-NMLZ-3SG book
'the book which the student bought'

c. *[cumhurbaşkanı-nın gel-diğ-i] haber-i*
president-GEN come-NMLZ-3SG news-3SG
'the news that the president has come'

The general noun-modifying construction may involve a participial marker (as in Karachay-Balkar), a general subordinator (as in Warndarang (Maran; Northern Territory, Australia)), a multifunctional complementizer (as in Chalcatongo Mixtec (Oto-Manguean; Mexico)), or even a finite clause with no overt subordinator (as in Japanese). In a number of cases, however, there is a gap strategy specific to relativization, as in Maale (Omotic; Ethiopia), exemplified in (17).

(17) Maale (Amha 2001: 160)

ʔííní [[ziginó mukk-é] ʔatsi]
3SG.M.NOM yesterday come-PFV.REL person.M.ABS
zag-é-ne
see-PFV-AFF.DECL
'He saw the man who came yesterday.'

Here the relative clause precedes the head noun and it contains no pronominal element coreferential to the relativized noun. The relative clause in Maale can be regarded as a specifically relative construction because it differs from other subordinate clauses in having no affix indicating the dependent status of the clause. And it also differs from independent sentences: whereas independent sentences are characterized by clause-final illocutionary force morphemes which classify the utterance as an assertion, interrogation, manipulative, etc., as in (18), the (restrictive) relative clause ends in one of the aspect/polarity suffixes *-é-*, *-á-*, *-uwá-*, or *-ibá-*, as in (17), and cannot be marked by the illocutionary force morphemes so that it cannot form a complete utterance on its own.

(18) Maale (Amha 2001: 160)

ʔatsí ziginó mukk-é-ne
person.M.NOM yesterday come-PFV-AFF.DECL
'The man came yesterday.'

2.2 Map 123 Relativization on obliques.

1.	Relative pronoun strategy	13
2.	Nonreduction strategy	14
3.	Pronoun retention strategy	20
4.	Gap strategy	55
5.	Not possible	10
	total	112

We distinguish five major groups of languages here. First, there are a number of languages that relativize upon obliques by employing the **relative pronoun strategy**, as exemplified in the Russian sentence in (19).

(19) Russian

Ja poterjal nož, kotorym
I lose.PST knife.ACC which.INSTR
ja narezal xleb.
I cut.PST bread
'I lost the knife with which I cut the bread.'

Second, a number of languages use a **nonreduction strategy** for relativizing on obliques, where the head noun appears as a full-fledged noun phrase within the relative clause, with the same three major subtypes, (i) correlative, (ii) internally headed relative clause, and (iii) paratactic clauses, as with relativization on subjects. The sentences in (20), (21), and (22) illustrate these three subtypes, respectively.

(20) Hindi (Comrie 1998: 62)

Maim jis ādmī se bāt kar rahā
I.DIR which.SG.OBL man to talk do PROG.SG M
thā vah kal bhārat jāegā.
be.IMPF.SG.M that.DIR.SG tomorrow India go.FUT M.SG
'The man [to whom I was talking] will go to India tomorrow.'
(lit. 'Which man I was talking with, he will go to India tomorrow.')

(21) Maricopa (Gordon 1986: 261)

Bonnie va-s-ii uuyem-sh havshuu-k
Bonnie house-DEM-at go.NOM-SUBJ blue-REAL
'The house Bonnie went to is blue.'

(22) Gooniyandi (Bunaban; Australia; McGregor 1990: 438)

ginharndi yoowooloo jijaggiddaa-nhi
you.know man we.are.speaking-of.him
wambiggoowaari
he.is.going.inside
'The man who we're talking about is going inside.'

The third major relativizing strategy with obliques is the **pronoun-retention strategy**. This strategy is exemplified in the sentence in (23) from Persian.

(23) Persian (Comrie 1998: 63)

mardhâi [ke ketâbhâ-râ be ânhâ dâde bud-id]
men that books-ACC to them given were-2SG
'the men that you had given the books to'
(lit. 'the men that you had given the books to them')

The fourth group of languages employ the **gap strategy**, as exemplified in the Korean sentence in (24).

(24) Korean (Comrie 1989: 151)

[Hyənsik-i kɨ kä-lɨl ttäli-n] maktäki
Hyensik-NOM the dog-ACC beat-REL stick
'the stick with which Hyensik beat the dog'

The fifth value represented on Map 123 is **nonrelativizable**, and it stands for those languages where obliques cannot be relativized upon directly. In such languages, the translation equivalent of relativizing upon an oblique in other languages is typically expressed by advancing the noun phrase in question to a position that can be relativized upon, e.g. by the use of applicative and/or passive constructions (for details, see Comrie 1989: 156–63).

Note that the assignment of a particular feature value to a particular language does not mean that this feature value is the only one that has been attested in that particular language. It only means that this particular feature value is considered to be the most frequent, or the canonical one in nonmarked contexts.

3 Geographical distribution

Map 122 shows the following areal-typological configurations with respect to relativizing on subjects. In Europe, the relative pronoun strategy predominates (see also Lehmann 1984: 109; Comrie 1998: 6; Haspelmath 2001: 1496–7). Note that this strategy stands out as being typically European since it is not found in Indo-European languages spoken outside Europe, and is exceptional more generally outside Europe.

In East Asia and South-East Asia, the gap strategy is the most frequent one (see also Comrie 1998: 78). The nonreduction relativizing strategy is most frequently employed in the languages of the Americas.

Map 123 reveals distinct areal-typological patterns, too. The gap strategy is the dominant relativizing strategy for obliques in South-East Asia, the Pacific area, and Australia. The relative pronoun strategy is characteristic of relativizing on obliques in Europe,

whereas pronoun retention is the most frequent relativizing strategy with obliques in the languages of Africa.

4 Theoretical issues

According to the Accessibility Hierarchy of Relativization (subject > direct object > indirect object > possessor) proposed in Keenan and Comrie (1977), it is easier to relativize on subjects than it is to relativize on any of the other positions, easier to relativize on direct objects than indirect objects, etc. One of the generalizations that has been made regarding the accessibility hierarchy is that the pronoun-retention strategy is preferred at the lower end of the hierarchy. This finds substantial confirmation in the present study, in that several languages use pronoun retention for relativizing upon obliques but not for relativizing upon subjects. In fact, few languages use pronoun retention to relativize upon subjects, though Babungo illustrates precisely this possibility, as in (11), repeated as (25):

(25) Babungo (Schaub 1985: 34)

mə̀	*yè*	*wə́*	*ntɨ́ə*	*fáŋ*	*ŋwə́*	*sɨ́*	*sàŋ*	*ghə̂*
I	see.PFV	person	that	who	he	PST2	beat.PFV	you

'I have seen the man who has beaten you.'

Note that to relativize upon direct objects, Babungo uses either a gap or pronoun retention, with the gap being obligatory with a few verbs in the perfective aspect, as in (26); in this Babungo provides an exception to the accessibility hierarchy generalization on the distribution of pronoun retention.

(26) Babungo (Schaub 1985: 34)

a. optional gap strategy

mə̀	*yè*	*wěembwā*	*fáŋ*	*tǐi*
I	see.PFV	child	who	father
wī	*sɨ́*	*sǎŋ*	*(ŋwə̀)*	
his	PST2	beat.PFV	(him)	

'I have seen a child whom his father had beaten.'

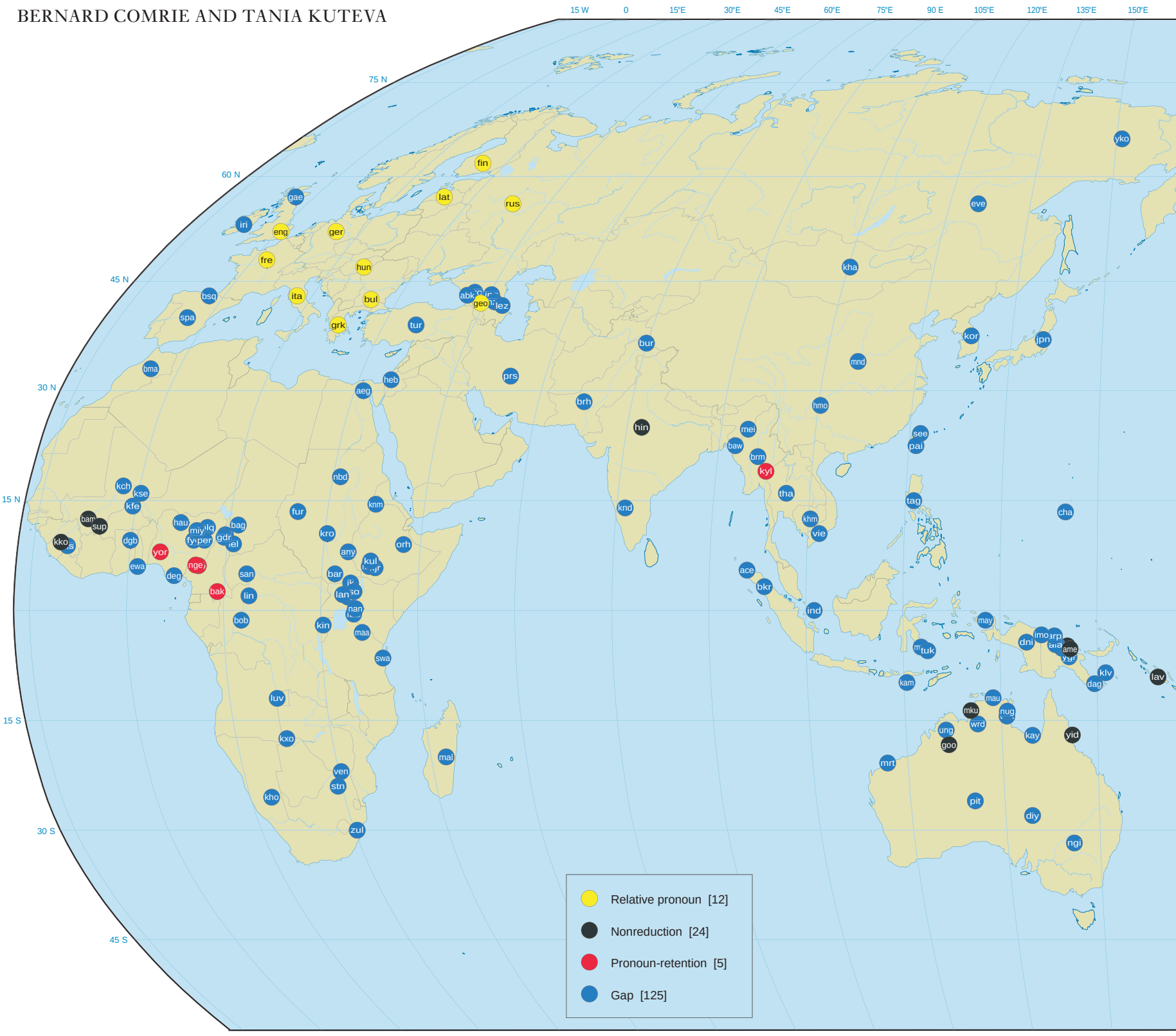
b. obligatory gap strategy

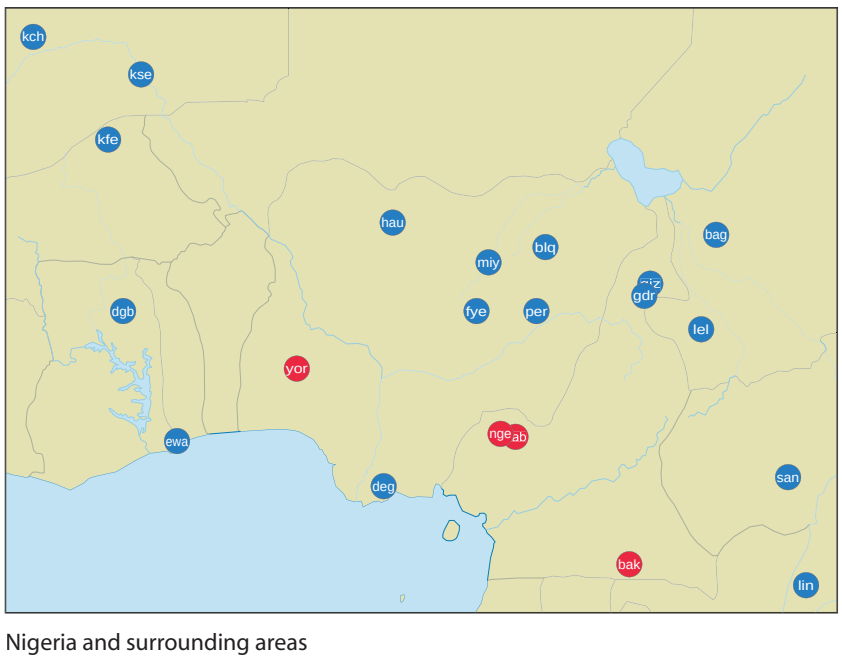
mə̀	*yè*	*ŋkáw*	*ŋkɨ́ə*	*fáŋ*	*Làmbí*	*kə̂*
I	see.PFV	chair	that	which	Lambi	give.PFV

'I have seen the chair which Lambi gave.'

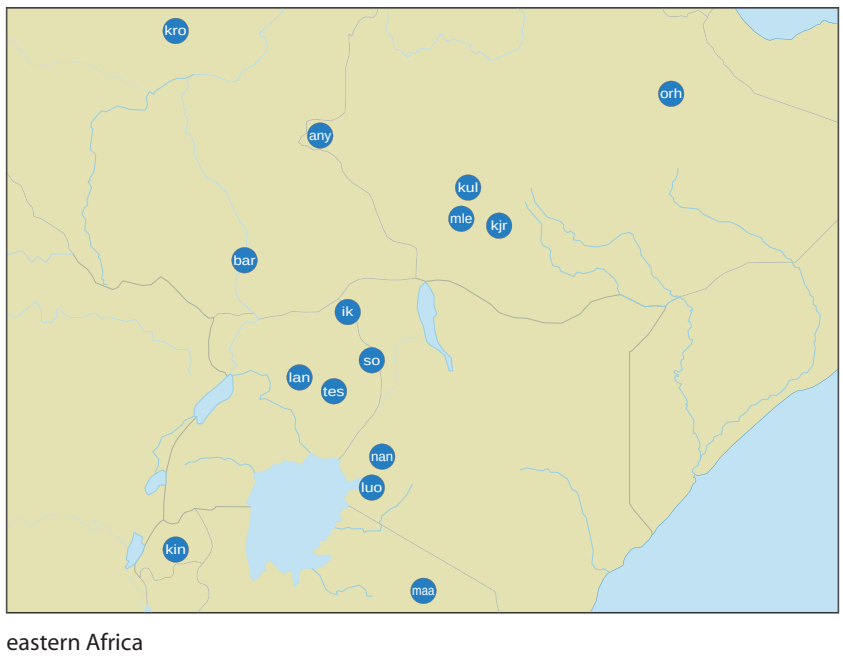
122 Relativization on Subjects

BERNARD COMRIE AND TANIA KUTEVA

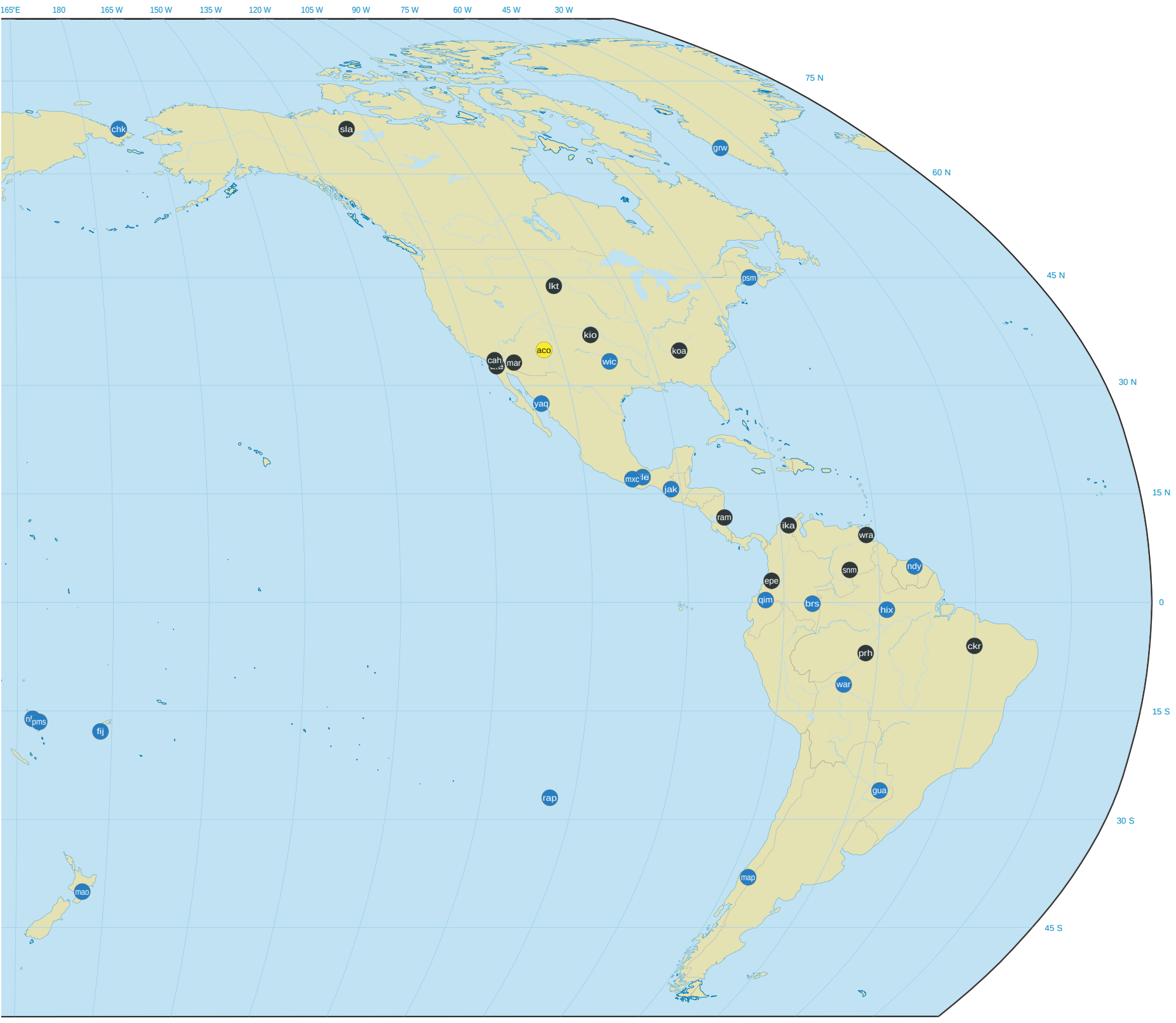


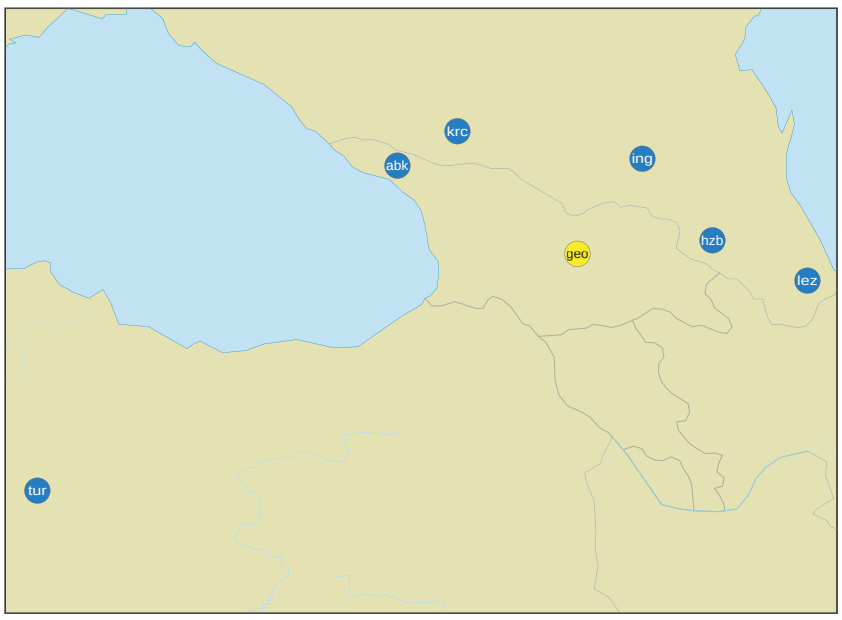


Nigeria and surrounding areas

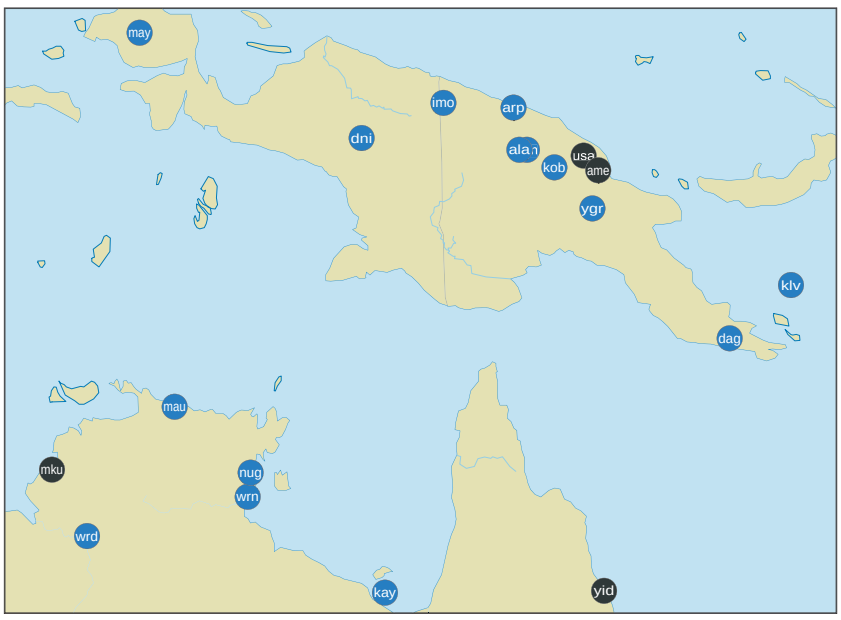


eastern Africa





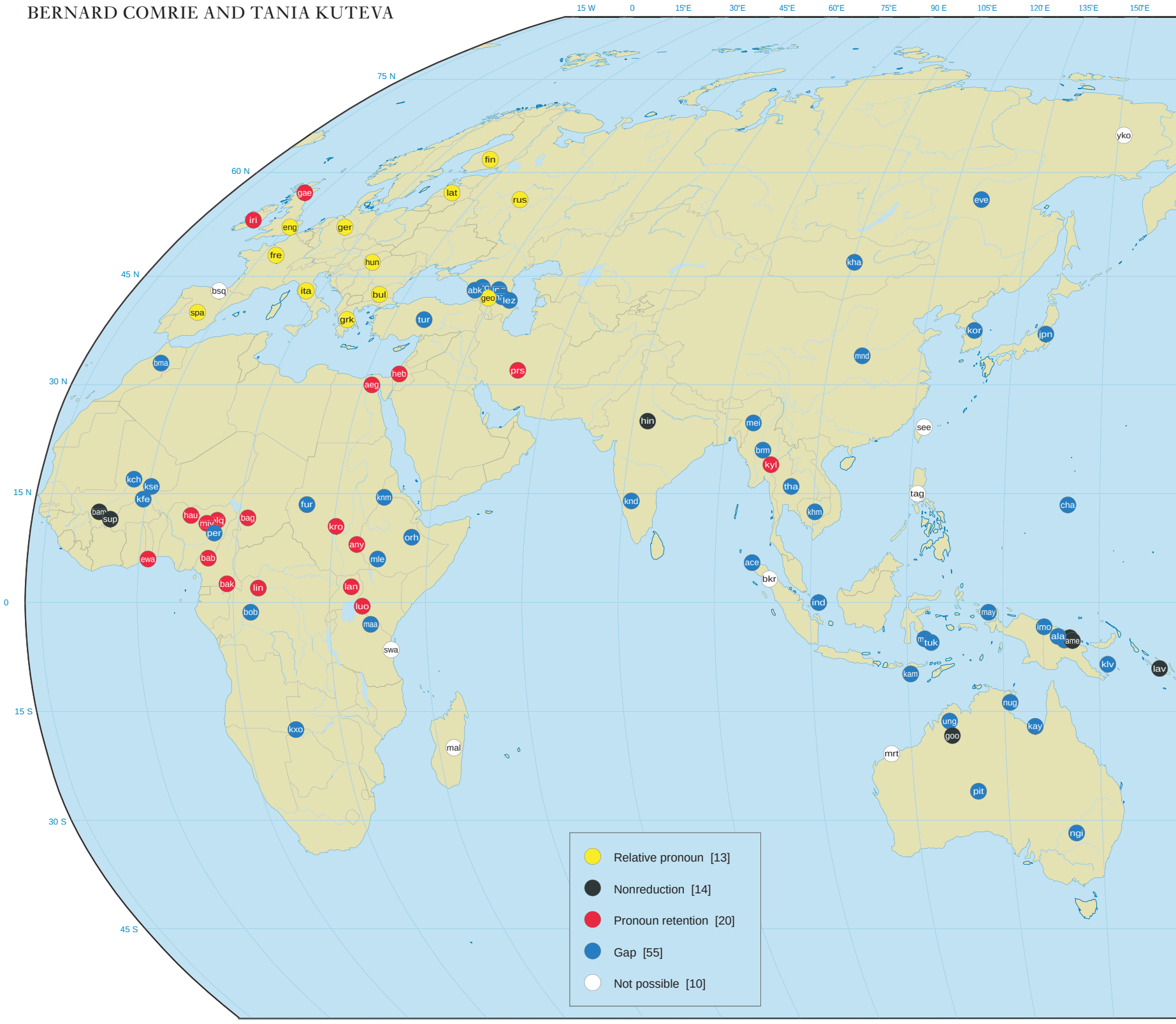
Caucasus and adjacent areas

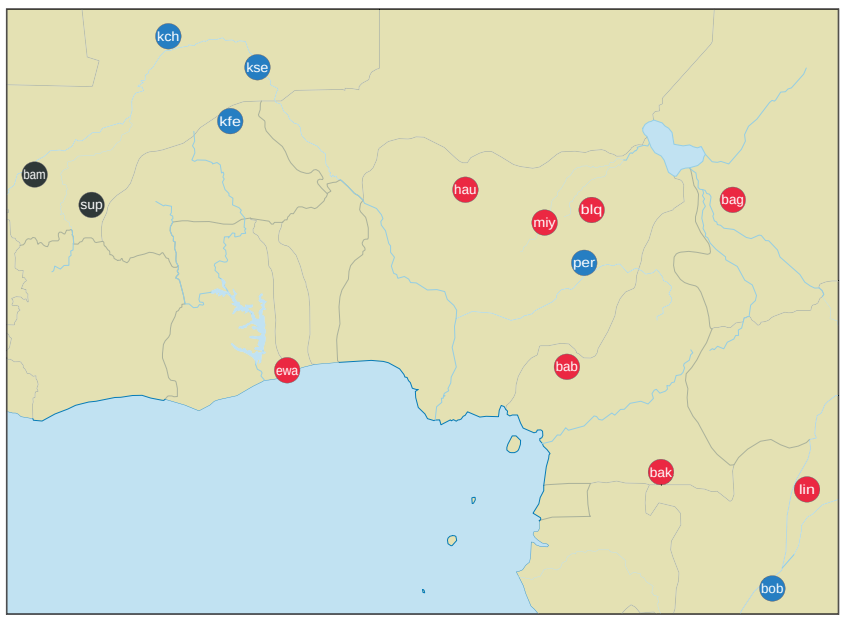


New Guinea and northern Australia

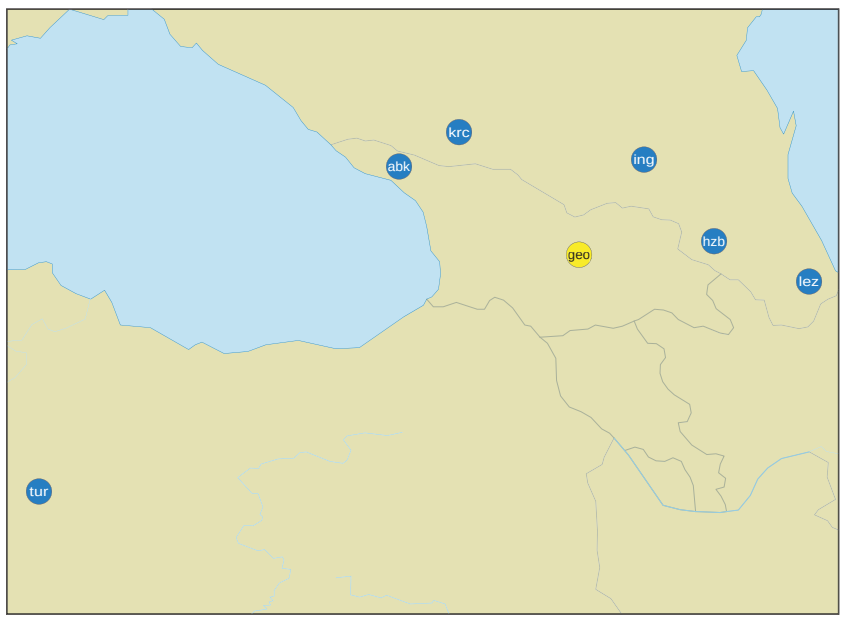
123 Relativization on Obliques

BERNARD COMRIE AND TANIA KUTEVA

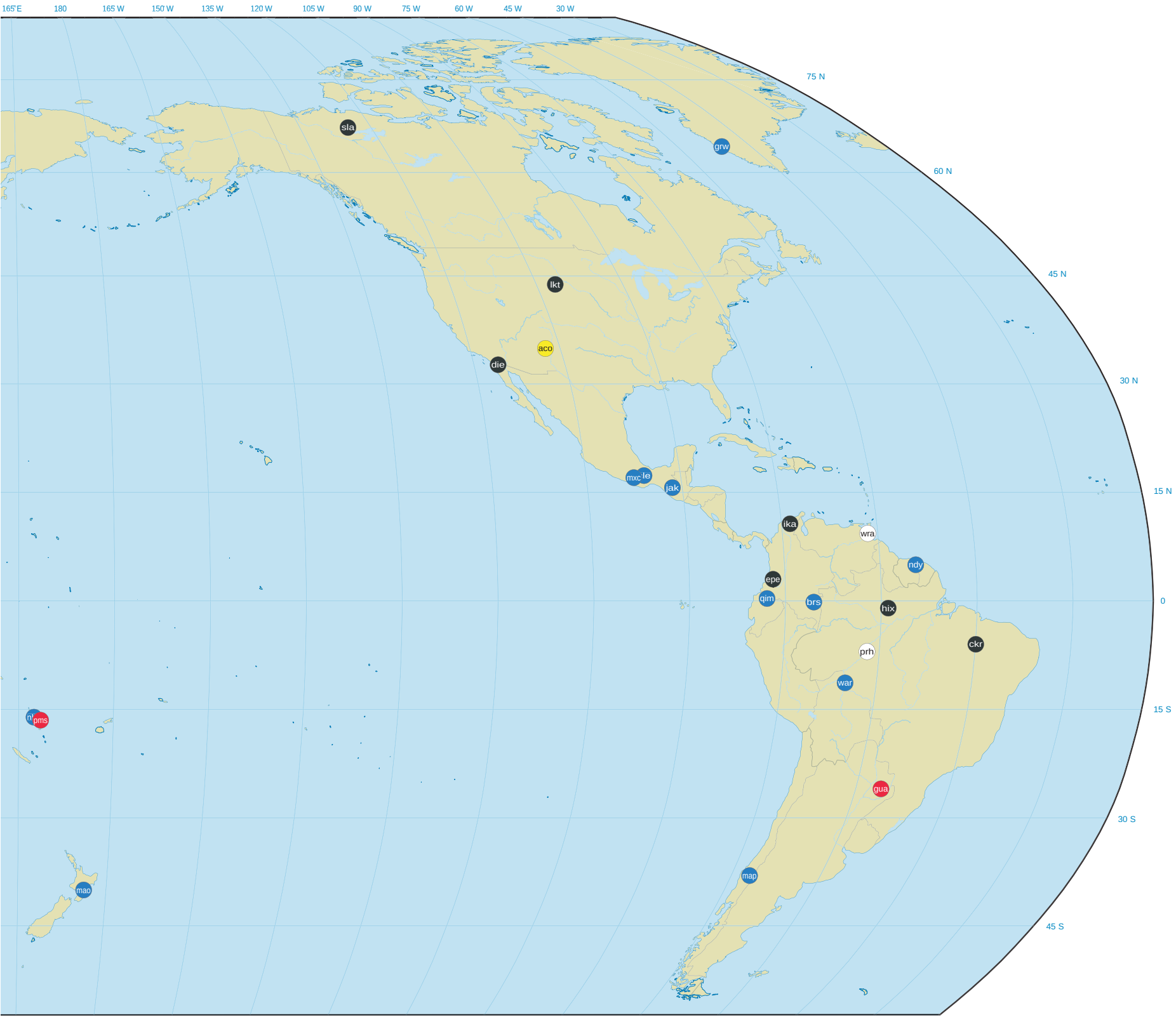


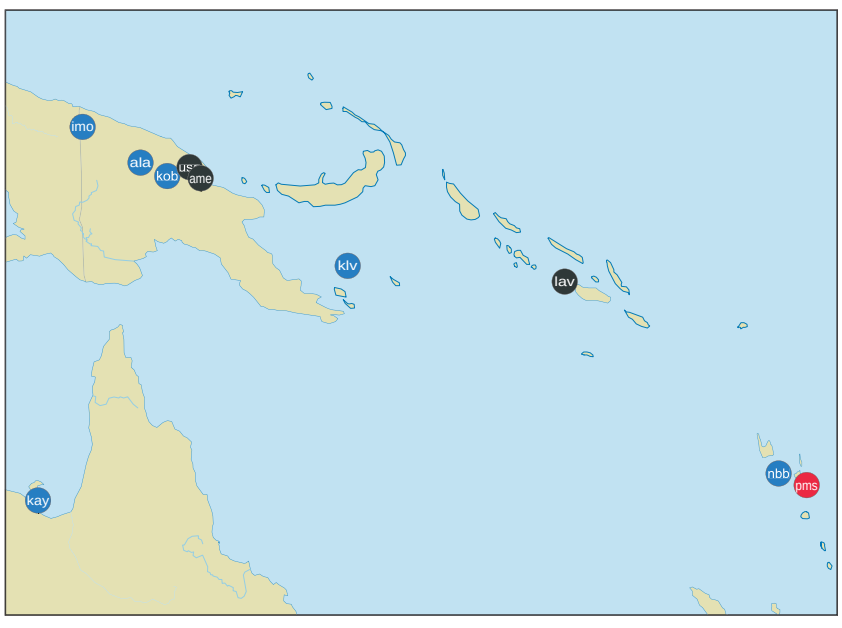


Nigeria and surrounding areas

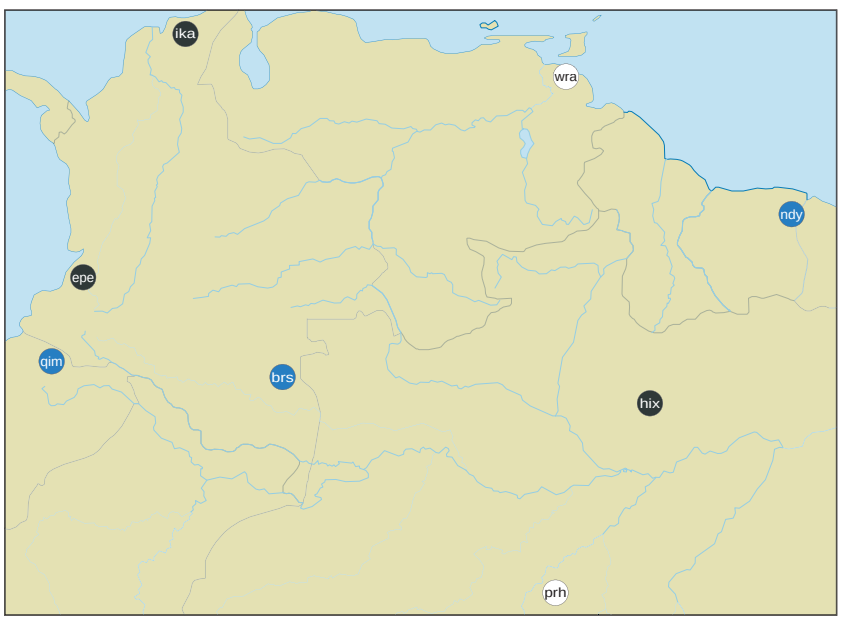


Caucasus and adjacent areas





Papua New Guinea and Melanesian Islands



northern South America

124 ‘Want’ Complement Clauses

MARTIN HASPELMATH

1 Overview of the values

This chapter looks at one important aspect of the syntax of ‘want’: whether the notional subject of its complement predication is **left implicit** when it is coreferential with the wanter, or whether it is **expressed overtly** in the complement clause by means of a pronominal element.

The first is the case in many languages which have special verb forms for such purposes (often called **infinitives**). The examples in (1) are from Turkish and Zulu, and the English translations also illustrate this type. (For the sake of clarity, the complement phrase or clause is enclosed in square brackets in all examples of this chapter.)

(1) a. Turkish (Wendt 1972: 56)
[Ankara-ya git-mek] ist-iyor-um.
Ankara-ALL go-INF want-IMPF-1SG
‘I want to go to Ankara.’

b. Zulu (Ziervogel et al. 1981: 210)
Ngi-funa [uku-dlala na-ye].
1SG-want INF-play with-him
‘I want to play with him.’

Notionally, there are two predications in these constructions: the predication of wanting, and the predication that expresses the desideratum. The latter is expressed in these examples by a complement phrase or clause that lacks an overt subject. The constructional meaning makes it clear that the notional complement subject is to be understood as being the same as the experiencer argument of ‘want’ in the main clause.

Alternatively, the complement may look more like a complete clause in that the complement subject is **expressed overtly** by a pronominal element, generally a simple verbal person-number marker. Verb forms used in such complement clauses are sometimes called **subjunctive**. The examples in (2a–b) are from Modern Greek and Egyptian Arabic.

(2) a. Modern Greek
Thél-o [na pá-o s-tin Athína].
want-1SG COMP go-1SG to-ART Athens
‘I want to go to Athens.’

b. Egyptian Arabic (Mitchell 1962: 144)
ʔana ʕaawiz [a-kallim il-mudiir].
I want 1SG-talk ART-manager
‘I want to talk to the manager.’

Literally, the sentence in (2a) can be translated ‘I want that I go to Athens’. The suffix *-o* in the subordinate verb *pá-o* expresses the subject in much the same way as the pronoun *I* does in English. When the wanter is a full noun phrase, it is of course not fully repeated in the complement clause, but is represented only by the appropriate pronominal element (e.g. Modern Greek *I Rúla thél-i na pá-i* [ART Roula want-3SG COMP go-3SG] ‘Roula wants to go’, literally ‘Roula wants that she go’).

Another possibility is that a language has **both constructions**, one that has an implicit complement subject, and another one that has an overtly expressed subject. An example of this type is Obolo (Delta Cross, Niger-Congo; southern Nigeria):

(3) Obolo (Faraclas 1984: 104, 112)
a. *M´-wèèk ígêgè íkpá.*
1SG-want INF.write letter
‘I want to write a letter.’

b. *M´-wèèk ǹ-gê íkpá.*
1SG-want 1SG-write letter
‘I want to write a letter.’

Finally, quite a few languages do not express the notion of wanting by means of an independent verb (or other lexeme), but by a **desiderative marker**. This can be a **verbal affix**, as in (4) from Apurinã (Arawakan; Amazonas, Brazil).

(4) Apurinã (Facundes 2000: 321)
nhi-nhika-ene-ta-ru
1SG-eat-DESID-VBLZ-3M.O
‘I wanted to eat it’

Alternatively, the desiderative marker can be an **uninflected particle** meaning ‘want’ that co-occurs with a fully inflected verb which expresses the desideratum predication. An example comes from Itzaj (Mayan; Guatemala).

(5) Itzaj (Hofling 2000: 48)
Tak u-wen-el.
DESID 3ABS-sleep-INCOMPL
‘S/he wants to sleep.’

With desiderative markers, the question of whether the complement subject is implicit or expressed does not arise because there is no complement phrase or clause. The two notional predications of wanting and eating/sleeping are expressed in a single clause in (4–5), and obviously the subject is expressed only once here.

Thus, there are five different language types that are distinguished on the map.

●	1. The complement subject is left implicit.	144
●	2. The complement subject is expressed overtly.	72
●	3. Both construction types exist.	14
●	4. ‘Want’ is expressed as a desiderative verbal affix.	45
◆	5. ‘Want’ is expressed as an uninflected desiderative particle.	8
	total	283

It should be noted that this chapter only covers constructions involving coreference between the wanter and the subject of the desideratum predication. When these two are not coreferential (as in ‘I want Roula to go to Athens’), many languages use a completely different construction.

2 Implicit complement subject

When the complement subject is left implicit, languages often use a special verb form such as the “infinitives” that were illustrated in (1) and that we also find in many Indo-European languages. Forms called “infinitives” generally lack not only subject-agreement marking, but also tense marking, and often they have nominal properties such as the possibility of occurring with adpositions. However, some languages do use verb forms that include a tense marker in implicit-subject complements, for instance Panyjima (Pama-Nyungan; Western Australia).

(6) Panyjima (Dench 1991: 200)
Ngatha purlpi [puntha-rta].
I.NOM want swim-FUT
‘I want to swim.’

And sometimes the bare stem of the verb is used in such constructions, even in languages with a fair amount of morphology. Example (7) is from Páez (Paezan; Colombia).

(7) Páez (Jung 1989: 141)
[Wakar-ti's-a'] *nasa* *[wala* *uy]*
white-DAT-TOP Páez much see
wê-xe-ts-me:-ta'.
want-IMPF-PROG-NEG-DECL.3PL
‘The Páez don’t want to see the whites much.’

In languages with little morphology, this is the most common pattern; see example (8) from Maori, which in addition to the verb stem uses a complementizer (this is also the case in English).

(8) Maori (Bauer 1993: 42)
Ka *piirangi* *ia* *[kia* *moohio* *ki* *te* *kootiro* *raa].*
TNS want he COMP know CASE ART girl that
‘He wanted to know that girl.’

3 Overtly expressed complement subject

In the great majority of cases, an overtly expressed subject in a ‘want’ complement clause is a person-number affix, like *-o* in Modern Greek (example 2a) and *(ʔ)a-* in Egyptian Arabic (example 2b). Languages with obligatory subject pronouns (i.e. non-affixal elements; see Chapter 101) generally leave the subjects of their ‘want’ complements implicit, but some exceptions to this generalization can be found, for instance in Dagbani (Gur, Niger-Congo; Ghana).

(9) Dagbani (Olawsky 1999: 35)
Doo *maa* *daa* *bɔri* *[ni* *o* *ŋubi* *noo].*
man the TD want SUBORD he chew fowl
‘The man wanted to eat fowl.’

In quite a few languages, the complement verb is additionally characterized by a special marker on the verb (often called **subjunctive** or a similar label), e.g. in Abkhaz (North-West Caucasian; Georgia).

(10) Abkhaz (Hewitt 1979: 33)
[S-abà-co *ħʷa* *eylə-y-k'aa-rc]* *yə-taxə̀-n.*
I-where-go say PREV-3SG.SUBJ-learn-PURP he-want-FIN
‘He wanted to learn where I was going.’

Sometimes a language has only very partial expression of the subject in the ‘want’ complement. Where the partial expression includes person information, it was counted as subject expression. In Tetun (Central Malayo-Polynesian; Timor), verbs have a person-number affix only for the first person singular. This inflection also occurs in ‘want’ complements, as in (11), so I have classified Tetun as a language of the second type, although in main clauses, given the impoverished agreement, independent pronouns are normally required in order to express the subject.

(11) Tetun (van Klinken 1999: 291)
Ha'u *la* *k-o'i* *[k-ola* *ó].*
I NEG 1SG-want 1SG-take you
‘I don’t want to take (i.e. marry) you.’

By contrast, in Hunzib (Daghestanian; eastern Caucasus), the complement verb only agrees with its intransitive subject in gender and number; there is no person indexing at all, and no agreement with the transitive subject. Hunzib was therefore classified as an implicit-subject language.

In quite a few languages, there is no unique word meaning only ‘want’, but ‘want’ is expressed by a verb that also means ‘say’ and/or ‘think’. In some of these cases, one wonders whether the ‘want’ meaning should not be ascribed to the mood inflection rather than to the main verb, for instance in Maricopa (Yuman; Arizona), where the complement verb takes the optative suffix *-ly*:

(12) Maricopa (Gordon 1986: 249)
[Kʷsede-sh *m-do-ly]* *m-aly'ii-m?*
doctor-SUBJ 2-be-OPT 2-think-Q.REALIS
‘Do you want to be a doctor?’

In a few North American languages, the ‘want’ complement behaves like a **direct-speech complement** in that the deictic centre is the wanter, so that the complement clause has a first-person subject. This is the case in Slave (Athapaskan; northwestern Canada).

(13) Slave (Rice 1989: 1235)
John *[náohzéh]* *enı̨dhę.*
John 1SG.OPT.hunt 3.want
‘John wants to hunt.’
(Literally: “John wants: ‘Let me hunt’.”)

Such languages, too, were classified as belonging to the explicit-subject type.

4 Both constructions (implicit subject and expressed subject) exist

In many cases, the two constructions that exist side by side in languages of this type are quite different in that, for instance, the implicit-subject construction requires a special infinitival form of the verb, as in (3a) from Obolo. But the two may also simply differ by the presence or absence of a subject pronoun. Ju|’hoan (Northern Khoisan; Botswana) comes close to this; in addition, the explicit-subject construction also requires the complementizer *ká*.

(14) Ju|’hoan (Dickens n. d.: 36)
Mí *kàrè* *(ká* *mí)* *g!à'ámá* *skóré.*
I want that I enter school
‘I want to enter the school.’

5 Desiderative markers

In most cases, the desiderative verbal affix only means ‘want’ (as in (4) from Apurinã). However, languages often use more general **nonrealized tense-mood forms** (called “irrealis”, “future”, etc.) where English would use *want*, especially when the reference is to a first-person subject. In general, such general tense-mood forms were not regarded as desideratives, but exceptions were made when the grammar author explicitly says that this mood form is the only way to express ‘want’, or when such a form is systematically and commonly rendered as ‘want’, even with third-person subjects, as in Tugun (Central Malayo-Polynesian; Maluku, Indonesia).

(15) Tugun (Douglas 1991: 86)
Marr-ala *la* *hira.*
3PL.IRR-take for them
‘They wanted to take them for themselves.’

When a language has several different construction types and one of them involves a desiderative marker, it was classified as belonging to the desiderative type, because this is the most grammaticalized construction type.

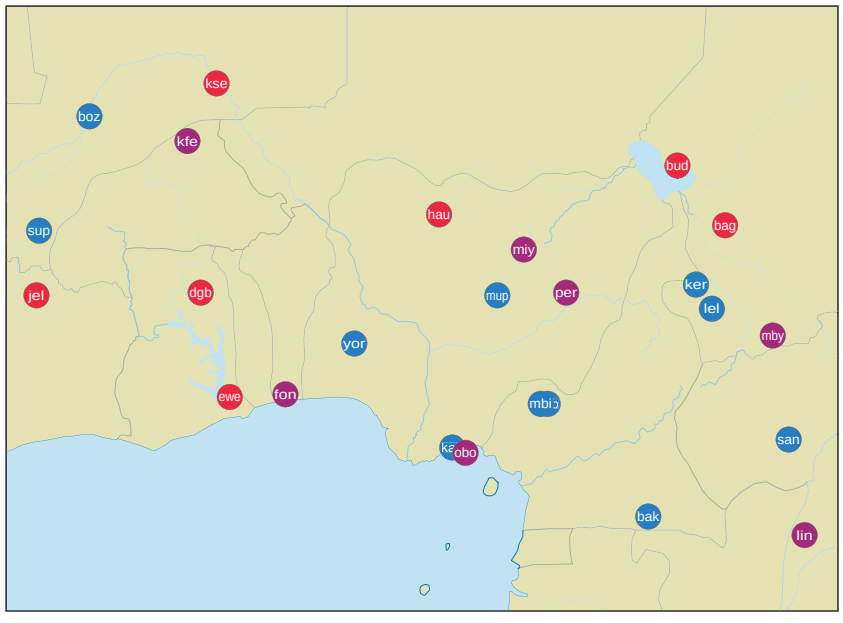
6 Geographical distribution

Only the explicit-subject type reveals an interesting geographical distribution: it shows a concentration in North America and Mesoamerica, as well as (remarkably, given the small amount of morphology in these languages) in West Africa. Within Europe, the explicit-subject Balkan languages stand out, but they fit well into an eastern Mediterranean/Middle Eastern context.

124 ‘Want’ Complement Clauses

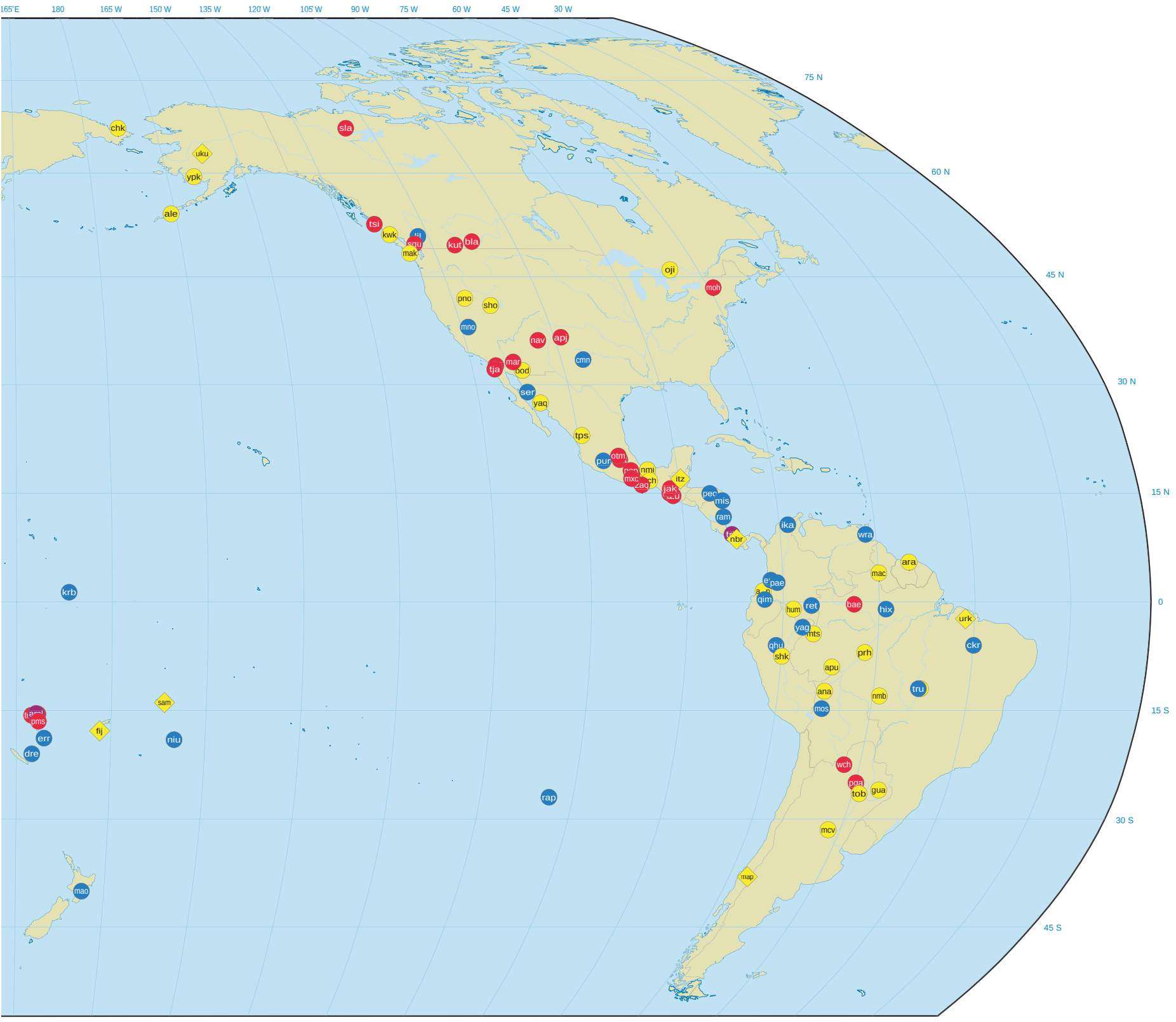
MARTIN HASPELMATH

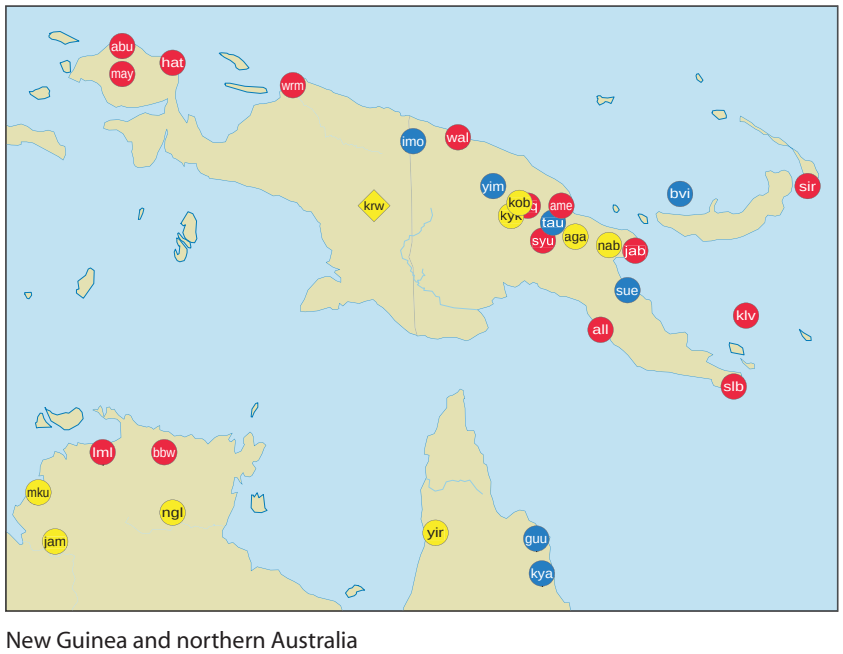
Subject is left implicit [144]
Subject is expressed overtly [72]
Both construction types exist [14]
Desiderative verbal affix [45]
Desiderative particle [8]



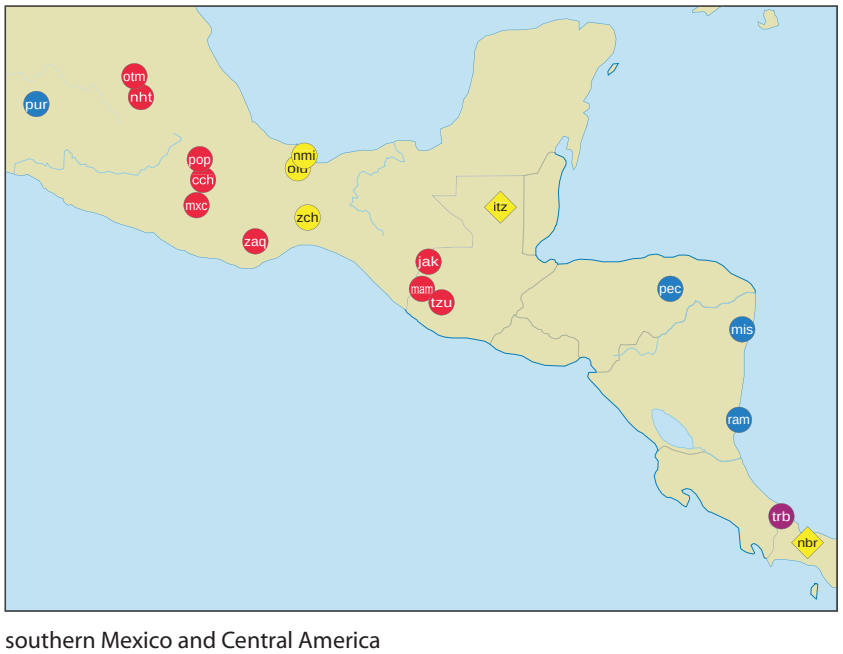
Nigeria and surrounding areas

areas surrounding the Black Sea





New Guinea and northern Australia



southern Mexico and Central America

125 Purpose Clauses

SONIA CRISTOFARO

1 Defining purpose clauses

This chapter investigates the form of the verb in purpose clauses. Following a common practice in typological research (see among others Stassen 1985 and Croft 1990: ch. 1), purpose clauses are defined in functional rather than morphosyntactic terms, as the latter turn out to be of restricted applicability in cross-linguistic comparison. A purpose construction is regarded as one encoding a particular relation between events. This relation is such that one of the linked events (the one coded by the main clause, or the main event) is performed with the goal of obtaining the realization of another one (the one coded by the purpose clause, or the dependent event). Typical cases of purpose relations involve motion predicates in the main event, as in (1). In this case, main and dependent event share a participant. However, non-motion predicates may also be involved, in which case main and dependent event need not share a participant, as is shown by (2) (the purpose clause is indicated by square brackets in the examples):

(1) *I went downtown [to buy books].*

(2) *I printed out a copy of this chapter [in order for you to look at it].*

In Cristofaro (2003: chs. 2, 6), the theoretical implications of the proposed functional definition of purpose clauses are discussed at length, and some pragmatic criteria are provided to identify purpose clauses under this definition.

The proposed definition of purpose clauses encompasses all of the traditional cases of purpose clauses (nonfinite clauses, possibly introduced by specific purpose conjunctions, as exemplified in (1) and (2), as well as finite clauses introduced by specific purpose conjunctions), plus some other cases that might not be regarded as such under traditional morphosyntactic criteria. One such case is provided by the serial verb construction in Mandarin. This construction involves a number of juxtaposed verbs, and can be translated as a purpose sentence (as in (3a)) or a coordinate sentence (as in (3b)) depending on the context:

(3) Mandarin (Li and Thompson 1973: 98)

a. *nǐ guì-xialai [qiú Zháng-san]*
you kneel.down beg Zhang-san
'You knelt down in order to beg Zhang-san.'

b. *nǐ guì-xialai qiú Zháng-san*
you kneel.down beg Zhang-san
'You knelt down and begged Zhang-san.'

In (3a) begging Zhang-san is the goal of the person's kneeling, but the relevant clause does not show any of the morphosyntactic hallmarks that differentiate purpose clauses from main clauses in other languages. However, under the proposed functional definition of purpose clauses, this clause can be regarded as a purpose clause, because this is the way Mandarin expresses the conceptual situation associated with purpose clauses in other languages.

2 Balancing and deranking

Following a distinction introduced by Stassen (1985), verb forms in purpose clauses are classified as either **balanced** or **deranked**, and the corresponding clauses will be called balanced and deranked purpose clauses (for some differences between the notion of deranking adopted here and Stassen's original formulation, see Cristofaro 1998 and 2003: ch. 3). A balanced verb form is one that can occur in an independent declarative clause, such as for instance an indicative form. Example (3a) above provides another instance of a balanced verb form: the linked verbs show the same morphological form, and each of them could occur in an independent declarative clause.

A deranked verb form is one that is structurally different from those used in independent declarative clauses. Deranking may take a variety of forms, such as lack of the categorial distinctions normally relevant to verbs in the language (e.g. tense, aspect, mood, or person agreement distinctions), or use of special elements not relevant to verbs in independent declarative clauses, e.g. nominalizers, case marking, or adpositions.

In example (4), from Lango (Nilotic; Uganda), the relevant verb form bears an infinitive suffix, has no aspect, mood, or person affixes, and is accompanied by the adposition *mê*:

(4) Lango (Noonan 1992: 245)
*lócà òbínô **[mê càmmò** gwènò]*
man 3SG.come.PERF for eat.INF chicken
'The man came to eat chicken.'

Deranked verb forms may also display special tense, aspect, mood, or person markers not used in independent declarative clauses. This is usually the case with subjunctives and the so-called dependent moods found in West Greenlandic (Fortescue 1984) or Abkhaz (Hewitt 1987). Example (5), from Italian, provides an instance of a subjunctive purpose clause. Tense, aspect, mood, and person are expressed by means of special affixes not used in independent declarative clauses (as is shown by the contrast between *scriv-e* 'write-3SG.IND' versus *accorg-**a*** 'realize-3SG.SUBJ'):

(5) Italian
Glie-lo scriv-e [affinché se ne
to.him/her-it write-3SG.IND in.order.for REFL of.it
***accorg-a**].*
realize-3SG.SBJV
'S/he writes that to him/her in order for him/her to realize that.'

In (6), from Finnish, person agreement on the deranked verb 'buy' is expressed by means of possessive suffixes, instead of the agreement suffixes found in independent declarative clauses:

(6) Finnish (Sulkala and Karjalainen 1992: 55)
*Men-en kauppa-an **[osta-a-kse-ni** ruoka-a].*
go-1SG shop-ILL buy-INF-TRANSL-1POSS food-PRTV
'I am going to the shop in order to buy food.'

Dependent moods and subjunctives display the same categorial distinctions as the verb forms used in independent declarative clauses (this is, for instance, the case in the Italian example in (5) above). Stassen (1985: 338–9) observes that in this case one might want to regard these forms as balanced (as is done for instance in Hengeveld 1998). However, since these forms cannot occur in independent declarative clauses, they will be regarded as deranked here.

It should be stressed that the distinction between balancing and deranking is not equivalent to that between **finiteness** and **nonfiniteness** (for which see Joseph 1983: ch. 2 and Koptjevskaja-Tamm 1994). The latter distinction is chiefly based on morphological criteria, i.e. nonfinite verb forms have reduced inflectional potential, while finite verb forms display full inflectional potential. The balancing/deranking distinction, on the other hand, is based exclusively on the ability of a verb form to occur in independent declarative clauses. Many verb forms with reduced inflectional potential (infinitives, participles, and the like, such as the verb *càmmò* in (4)) cannot occur in independent declarative clauses, and thus count as deranked forms. However, verb forms with reduced inflectional potential do count as balanced if they can occur in independent declarative clauses. This is, for instance, the case of nominalized verb forms in Kayardild, which are used in both purpose clauses and independent declarative clauses (where they express uncompleted action: Evans 1995: 470–4). On the other hand, some verb

forms with full inflectional potential, such as subjunctives or dependent moods, cannot occur in independent declarative clauses, and thus count as deranked.

An apparent problem for the balancing/deranking distinction is provided by verb forms bearing clitics or affixed conjunctions, such as the purpose conjunction *-ghoo* in the following Muna (Austronesian; Indonesia) example:

(7) Muna (van den Berg 1989: 264)
no-kala *we* *sangku* ***[nae-wei-ghoo]***
3SG.REAL-go LOC forest 3SG.IRR-clear-PURP
'He goes to the forest to clear (it).'

One might want to regard such forms as deranked, because the complex verb + clitic/affixed conjunction could not occur in an independent declarative clause. However, the verb as such often displays the same categorial distinctions as the forms used in independent declarative clauses. This is, for instance, the case in Muna, where the verb in the purpose clause is inflected for the same categories (person and realis) appropriate for verbs in independent declarative clauses. Therefore, these forms will be regarded here as balanced, unless some proper deranking strategy (e.g. lack of inflectional distinctions, expression of inflectional distinctions by means of special forms not used in independent clauses) is also at work.

For any given language, a purpose clause can be coded by deranked verb forms only, by balanced verb forms only, or by either deranked or balanced verb forms. Map 125 shows the geographical distribution of these three situations in the world's languages.

1. Balanced	38
2. Balanced/deranked	30
3. Deranked	102
total	170

3 Geographical distribution

The vast majority of the world's languages have deranked purpose clauses, as can be seen from the feature-value box at the end of the previous section. Languages with exclusively balanced purpose clauses are concentrated in South-East Asia and the Pacific islands. Languages with deranked purpose clauses are common all over the world, except for mainland South-East Asia and Indonesia.

Languages with exclusively deranked purpose clauses are dominant in the Americas, though languages with both deranked and balanced purpose clauses, or exclusively balanced purpose clauses, are also widespread there. Languages with exclusively balanced purpose clauses are relatively common in North and Central America, but rarer in South America. Languages with both deranked and balanced purpose clauses are more common in North and Central America than in South America.

Languages with exclusively deranked purpose clauses are also dominant in Africa. Languages with both deranked and balanced purpose clauses, or balanced purpose clauses only, are also relatively common, and the former are more frequent than the latter.

Europe is divided into two clear-cut areas, with languages with both balanced and deranked purpose clauses towards the north, and languages with exclusively deranked purpose clauses towards the south.

Most languages in central, southern, and eastern Asia display exclusively deranked purpose clauses.

Languages with exclusively deranked purpose clauses are dominant in Australia, although a few languages with exclusively balanced purpose clauses, or both balanced and deranked purpose clauses, are found in the north. In the Pacific islands, languages with exclusively deranked purpose clauses are as common as languages with exclusively balanced purpose clauses; languages with both deranked and balanced purpose clauses are marginal here (the only such language in the sample is Amele, in Papua New Guinea).

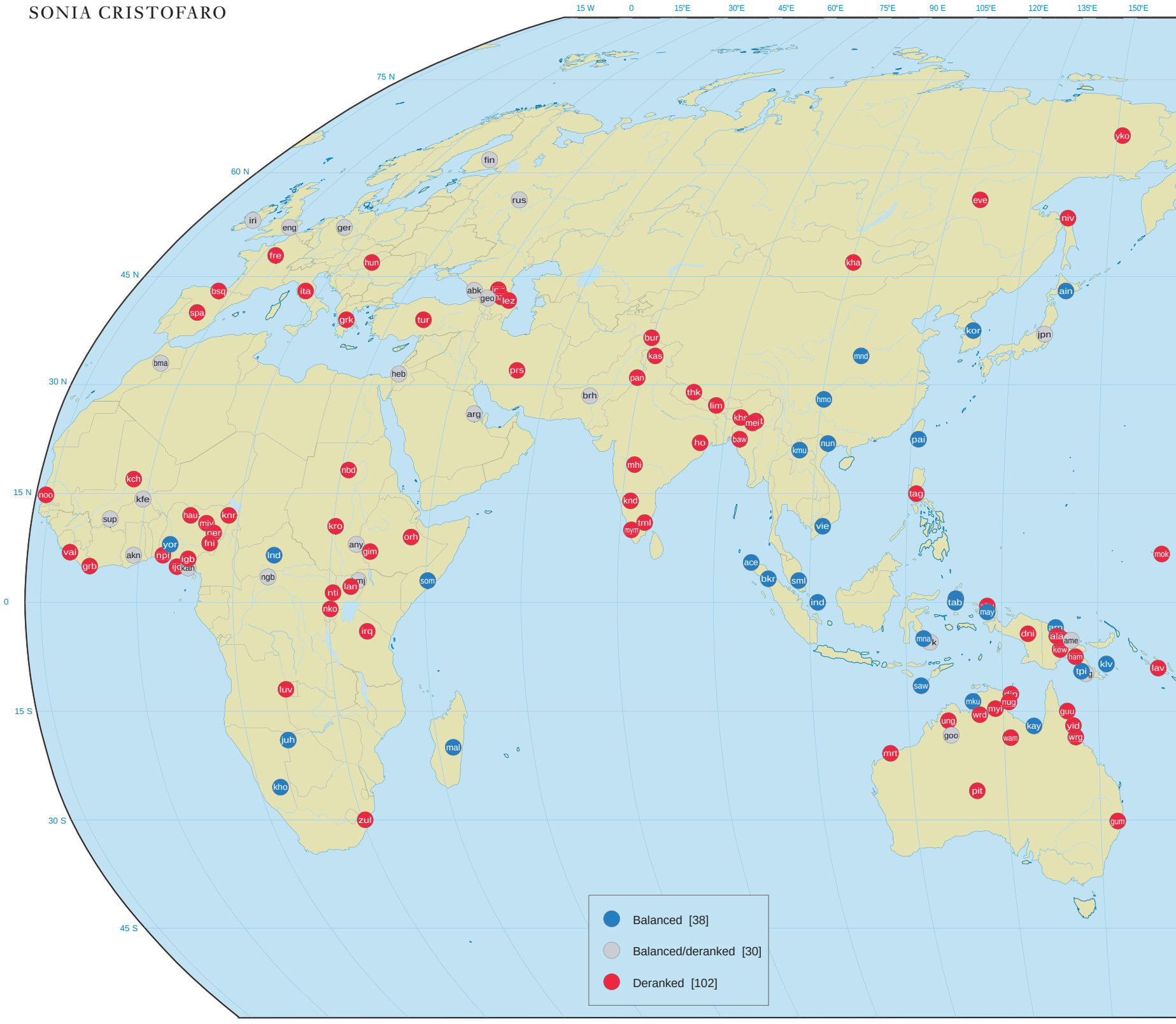
4 Theoretical issues

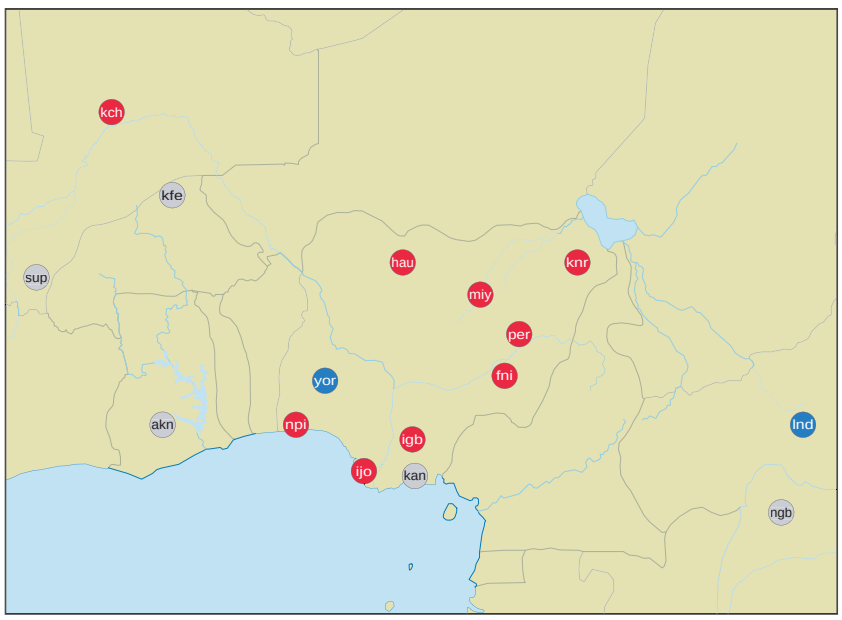
The most striking fact about the distribution of balanced and deranked purpose clauses across the world's languages is the dominance of languages with deranked purpose clauses. In Cristofaro (1998, 2003), it is argued that there is a functionally motivated connection between the structural features of deranked verb forms and the semantic features of purpose relations. In purpose relations, the time reference of the dependent event is predetermined as posterior to that of the main event, and the two events often share a participant (see example (1) above). Thus, the time reference of the dependent event, and the participant shared by main and dependent event, need not be expressed overtly. This leads to the use of verb forms not marked for time and person agreement, i.e. deranked forms. Thus, the use of deranked forms in purpose clauses reflects an **economic motivation**, i.e. the tendency not to express what is entailed or recoverable from the context (on this point see Haiman 1983 and 1985: 102–30). Furthermore, the dependent event in a purpose relation is unrealized, and unrealized events are often expressed by means of special verb forms not used for realized events, such as subjunctives and the like—these forms are also regarded as deranked here. Finally, purpose relations involve a high degree of semantic integration between main and dependent event (in the sense defined in Givón 1980 and 1990: ch. 13). Use of deranked verb forms often leads to syntactic integration between clauses, in that a number of categories relevant to the dependent verb (tense, aspect, mood, person) are not expressed, so that the information about these categories must be recovered from the main verb. Hence the use of deranked forms in purpose clauses reflects an **iconic motivation**, i.e. semantic integration between events is iconically reflected by syntactic integration between clauses.

The functional correspondence between the structural features of deranked verb forms and the semantic features of purpose relations explains why balanced purpose clauses seem to be disfavoured across the world's languages. In fact, as is shown in Cristofaro (1998, 2003), an implicational hierarchy exists such that if a language uses balanced verb forms for purpose clauses, than it uses balanced verb forms for the other adverbial clause types, as well as for relative clauses and most types of complement clause.

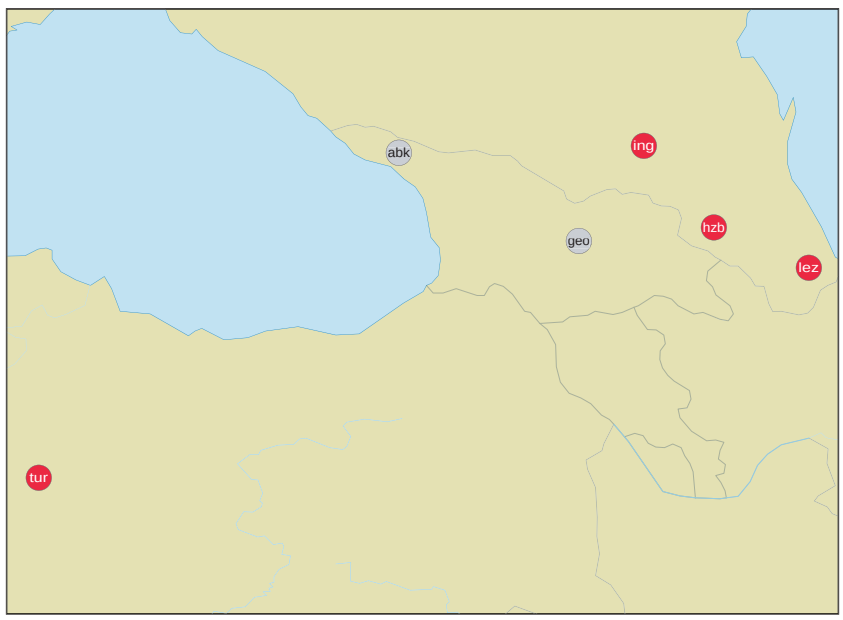
125 Purpose Clauses

SONIA CRISTOFARO

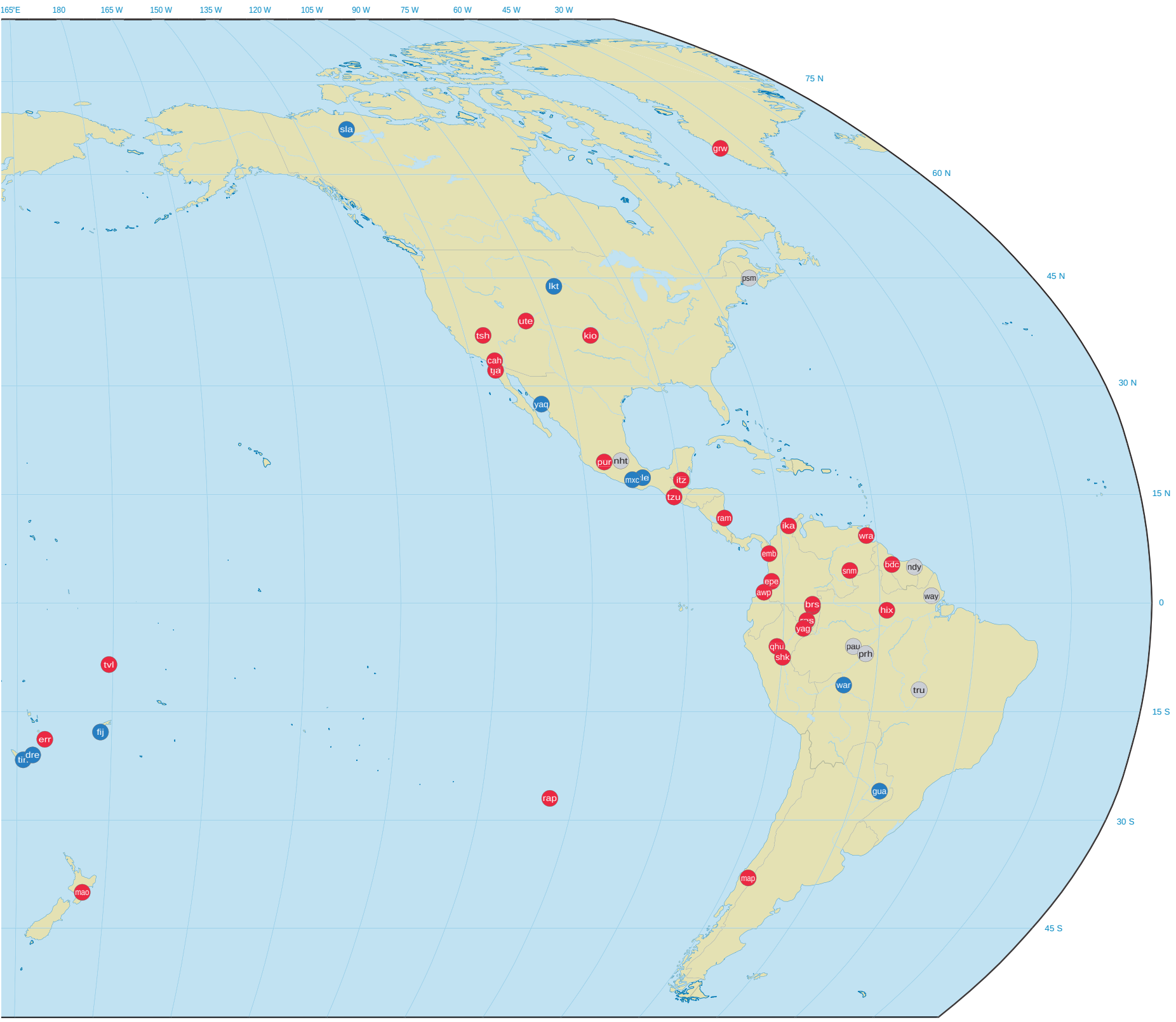


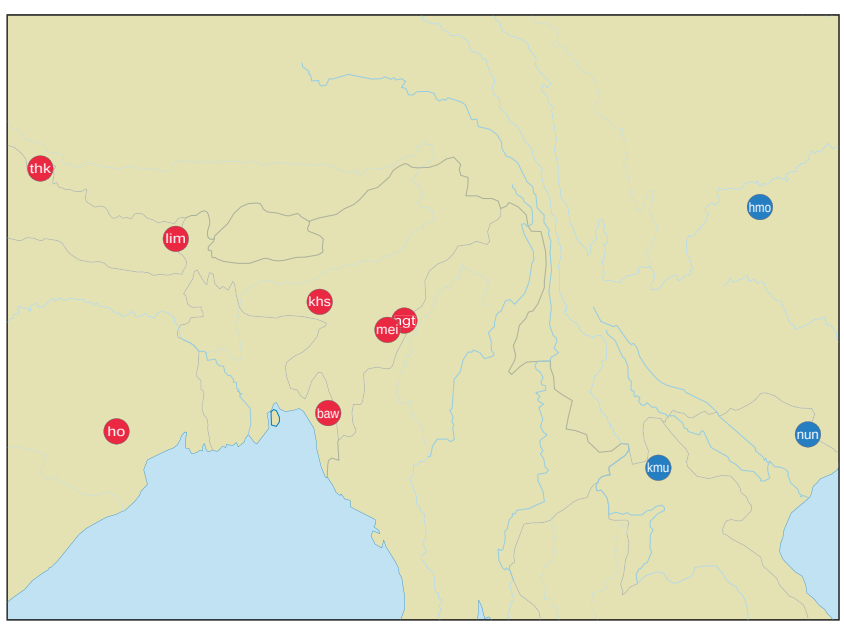


Nigeria and surrounding areas

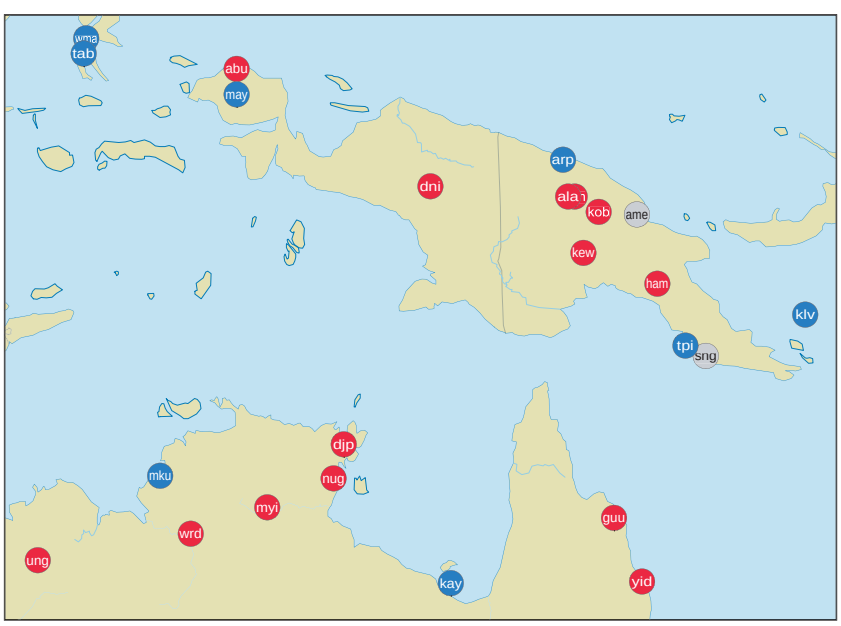


Caucasus and Black Sea area





from Nepal to Vietnam



New Guinea and northern Australia

126 'When' Clauses

SONIA CRISTOFARO

1 Defining the values

This chapter examines the form of the verb in 'when' clauses. Examples (1)–(3) provide instances of 'when' clauses in English (throughout the chapter, 'when' clauses will be enclosed in square brackets in the examples):

(1) *[When I go there], there is hardly anybody around.*

(2) *[When writing the next chapter], you should deal with these issues.*

(3) *[When the Nazis came to power], several intellectuals left Germany.*

As was done for purpose clauses (Chapter 125), 'when' clauses are defined in functional, rather than morphosyntactic, terms. A 'when' construction is regarded as one encoding a temporal relation between two events, such that there is a temporal overlap between the two. The exact extent of the overlap is unspecified and subject to variation. For instance, in (1) and (2) the linked events are simultaneous, while in (3) they may or may not be simultaneous—in fact, there might be an interval of some days, or even months or years, between them. The theoretical implications of this definition are discussed in Cristofaro (2003: chs. 2, 6), where a number of pragmatic criteria are provided to identify 'when' clauses under this definition.

The proposed functional definition of 'when' clauses encompasses both traditional cases of 'when' clauses (e.g. nonfinite 'when' clauses, as in (2), or finite clauses introduced by specific temporal conjunctions, as in (1) and (3)) and clauses that might not count as such under traditional morphosyntactic criteria. One such case is provided by Tzutujil (Mayan; Guatemala):

(4) Tzutujil (Dayley 1985: 367)
Kongáana q'ab'arik nb'ajni [k'o q'ojoom]
tremendous drinking is-done exist marimba
'Tremendous drinking is done on there being a marimba (when there is a marimba).'

In this case, the linked clauses are simply juxtaposed, and the verbs in each of them show the same morphological form. The linkage between the two is not signalled by any grammatical marker, and has to be inferred on contextual grounds. However, the clause referring to the *marimba* expresses the same kind of conceptual situation associated with 'when' clauses in other languages (this is proved by the application of the pragmatic tests discussed in Cristofaro 2003: ch. 2). Therefore, under the proposed functional definition of 'when' clauses, it can be regarded as a 'when' clause.

A similar case is found in Canela-Krahô (Macro-Ge; Brazil):

(5) Canela-Krahô (Popjes and Popjes 1986: 139)
[pê wa i-pỳm], pê inxê ty
PST 1 1-fall PST mother die
'My mother died when I was born.'

In this case, too, there is no marker signalling the linkage between the two clauses, and the verbs show the same morphological form. However, the clause expressing the speaker's being born is regarded as a 'when' clause here, because that is the way Canela-Krahô expresses the conceptual situation associated with 'when' clauses in other languages.

Verb forms in 'when' clauses may be either **balanced** or **deranked**, and the corresponding 'when' clauses will henceforth be indicated as balanced 'when' clauses and deranked 'when' clauses. The notions of balancing and deranking were originally introduced by Stassen (1985), and are defined in greater detail in Chapter 125. A balanced verb form is one that can occur in an independent declarative clause, e.g. an indicative verb form. Examples (1), (3), (4), and (5) above provide instances of balanced 'when' clauses.

A deranked verb form is one that cannot be used in independent declarative clauses. A deranked verb form may lack some or all of the categorial distinctions relevant to verbs in the language (such as tense, aspect, mood, or person agreement distinctions), or display special markers not used in independent clauses, e.g. special tense, aspect, mood, or person markers, nominalizers, case markers, or adpositions.

In example (6) from Maricopa (Yuman; Arizona), the verb is not inflected for tense, aspect, and mood, and bears the suffix *-haay* 'yet' as well as the locative case marker *ly* 'in':

(6) Maricopa (Gordon 1986: 270)
[']-ashvar-haay-ly] *'-nchen-sh iima-k*
1-sing-yet-in 1-old.sib-SBJV dance-REAL
'When I sang, my brother started to dance.'

In the Maori (Polynesian; New Zealand) example in (7), the relevant verb form is specified for mood (subjunctive), but not for tense or aspect.

(7) Maori (Bauer 1993: 62)
[Kia tae *mai te moni],* *maa.ku*
SBJV arrive hither the money INTGEN.1SG
e hoko mai he koha maa.u
T/A barter hither a gift INTGEN.2SG
'When the money arrives, I will buy you a present.'

Example (8), from Abkhaz (North-West Caucasian; Georgia), illustrates the use of so-called dependent moods. These are verb forms that cannot occur in independent declarative clauses, but that display all of the categorial distinctions relevant to verbs in the language. However, these distinctions are often realized by means of special forms, different from those used in independent declarative clauses. For instance, within the tense system of Abkhaz, each tense distinction can be realized by two different affixes (the so-called finite and nonfinite affixes) depending on whether the relevant verb form is used in independent or in dependent clauses (Hewitt 1987: 10–11). Thus in (8) both the main and the 'when' clause are in the aorist, but the verb in the 'when' clause has zero marking, while the main verb has the affix *–yt':*

(8) Abkhaz (Hewitt 1987: 138)
[d-anə̀-z-ba] *a-šʷqʷ'ə̀*
her-when-I-see.NONFIN.AOR the-book
(ø)-lə̀-s-ta-yt'
it-to.her-I-give-FIN.AOR
'When I saw her, I gave her the book.'

Stassen (1985: 338–9) observes that forms such as the Abkhaz nonfinite aorist might be regarded as balanced, because they display the same categorial distinctions as the forms used in independent declarative clauses. However, as was pointed out in Chapter 125, these forms will be regarded as deranked here, because they cannot themselves occur in independent declarative clauses.

The balancing/deranking distinction overlaps with, but is not equivalent to, the distinction between finiteness and nonfiniteness. For discussion of the relevant issues, see Chapter 125.

As was observed in Chapter 125, an apparent problem for the balancing/deranking distinction is represented by verb forms bearing clitics or affixed conjunctions, such as the clitic *-enean* in the following Basque example:

(9) Basque (Saltarelli et al. 1988: 43)
*[Etxe-ra irits-i n-in-tz-****enean]****,*
house-ALL.SG arrive-PERF 1.SG.ABS-PST-AUX-(PST)-when
kontura-tu n-in-tz-en gatza
realize-PERF 1SG.ABS-PST-AUX-(PST)-COMP salt (ABS.SG)

eros-te-a *ahaz-tu*
buy-NMLZ-ABS.SG forget-PERF
z-i-tza-ida-la
3.ABS-PST-AUX-1.DAT.SG-(PST)-COMP
‘When I arrived home, I realized that I had forgotten to buy salt.’

The complex “verb + clitic/affixed conjunction” could not occur in an independent declarative clause, and thus one might want to regard the relevant verb form as deranked. However, the verb as such often displays the same categorial distinctions as the forms used in independent declarative clauses (this is, for instance, the case in Basque); if so, the relevant verb form is regarded here as balanced.

For any given language, ‘when’ clauses can be coded by balanced verb forms only, by deranked verb forms only, or by either balanced or deranked verb forms. Map 126 shows the distribution of these three types in the world’s languages.

1.	Balanced	84
2.	Balanced/deranked	39
3.	Deranked	51
	total	174

2 Geographical distribution

Languages with exclusively balanced ‘when’ clauses are overwhelmingly dominant in mainland South-East Asia, Indonesia, and the Pacific islands. In fact, in mainland South-East Asia no languages with deranked ‘when’ clauses are found (the same pattern was observed for purpose clauses, and this reflects the implicational hierarchy for the coding of subordination presented in Cristofaro (1998, 2003); see Chapter 125). On the other hand, languages with exclusively balanced ‘when’ clauses are absent from western Europe, as well as from a large area extending from eastern Europe to Siberia in the east and India in the south. Languages with exclusively deranked ‘when’ clauses are absent from Europe, while languages with both deranked and balanced ‘when’ clauses are absent from North America.

In both North and South America, languages with exclusively deranked ‘when’ clauses are approximately as frequent as languages with exclusively balanced ‘when’ clauses. In Central America, on the other hand, languages with exclusively balanced ‘when’ clauses are dominant. Languages with both deranked and balanced ‘when’ clauses are marginal in both Central and South America.

Languages with exclusively balanced ‘when’ clauses are dominant in Africa, but languages with exclusively deranked ‘when’ clauses and languages with both deranked and balanced ‘when’ clauses are also frequent.

Most languages in Europe display both deranked and balanced ‘when’ clauses. On the other hand, languages with exclusively deranked ‘when’ clauses are dominant in central, southern, and eastern Asia. This pattern is similar to that found in the same area for purpose clauses (see Chapter 125), except for a pocket of languages with both deranked and balanced ‘when’ clauses in India.

In Australia, languages with exclusively balanced ‘when’ clauses are approximately as frequent as languages with exclusively deranked ‘when’ clauses, while languages with both deranked and balanced ‘when’ clauses are marginal.

3 Theoretical issues

Both balanced and deranked ‘when’ clauses are widespread across the world’s languages. Balanced ‘when’ clauses seem to be more common than deranked ‘when’ clauses, but it should be pointed out that the frequency difference between the two is not very significant. Also, there is a high concentration of languages with balanced ‘when’ clauses in some geographical areas (mainland South-East Asia, Indonesia, and Oceania), which suggests that the frequency patterns for balanced clauses might be areally and or genetically biased.

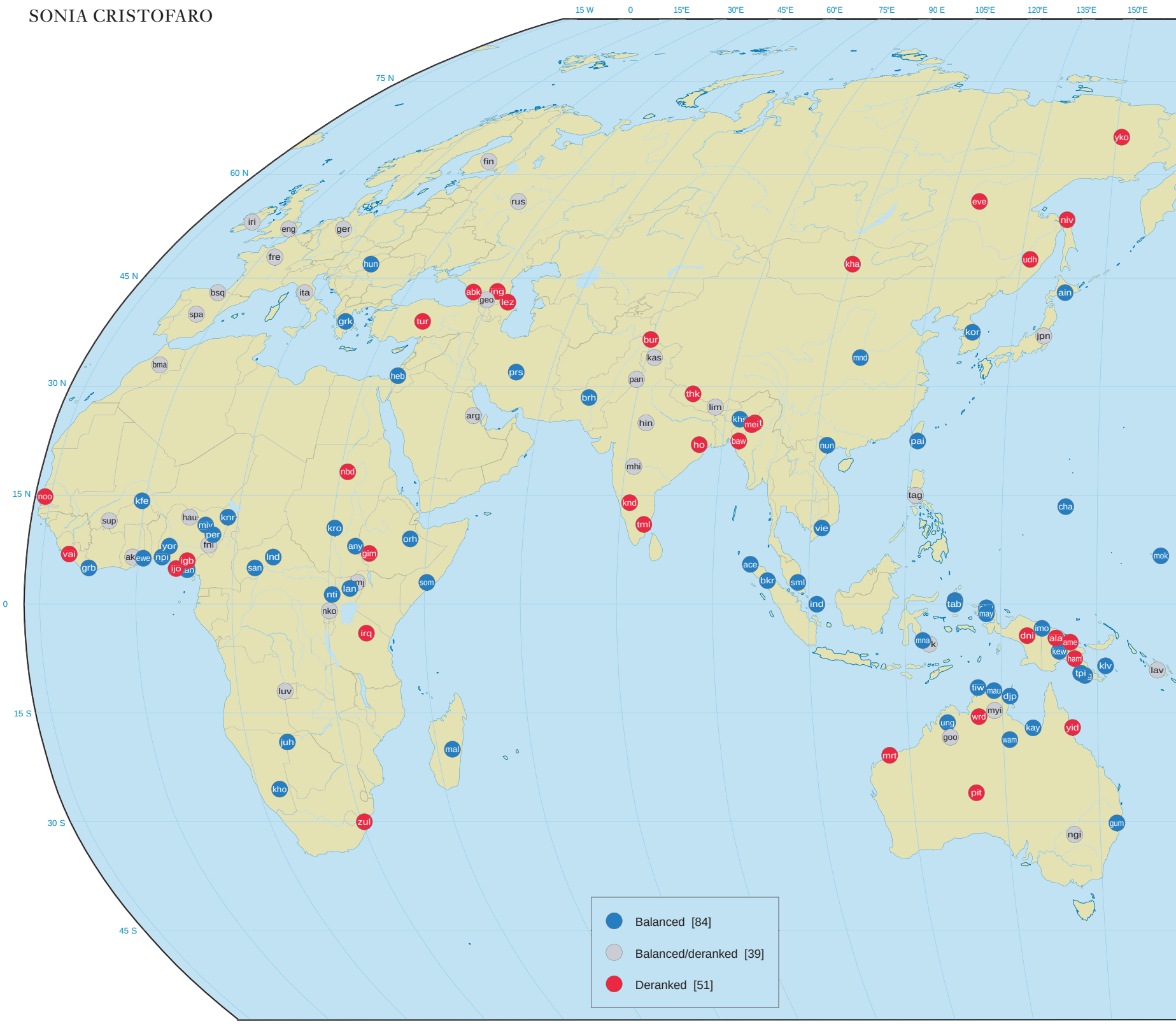
In Cristofaro (1998, 2003) it is shown that there is a functionally motivated correlation between the semantic features of ‘when’ relations and some of the structural features of deranked verb forms. ‘When’ relations imply that the dependent event takes place at the same time as the main event (though see the observations at the beginning of §1). This means that there is no need to overtly express the time reference and the mood value of the dependent event in the ‘when’ clause, as these can be recovered from the context. Deranked verb forms often do not express tense and mood distinctions, so their use in ‘when’ clauses reflects an **economic motivation**: the tendency not to express what is entailed or recoverable from the context (Haiman 1983, 1985; the same motivation holds for the use of deranked verb forms in purpose clauses; see Chapter 125). Also, ‘when’ relations have no implication as to whether or not the linked events have already taken place at the speech-act moment. If the two events are unrealized at the speech-act moment (as is the case in (1) and (2)), this may be indicated by means of special verb forms not used for realized events, as is the case in (7). These forms (conditionals, subjunctives, and the like) count as deranked verb forms.

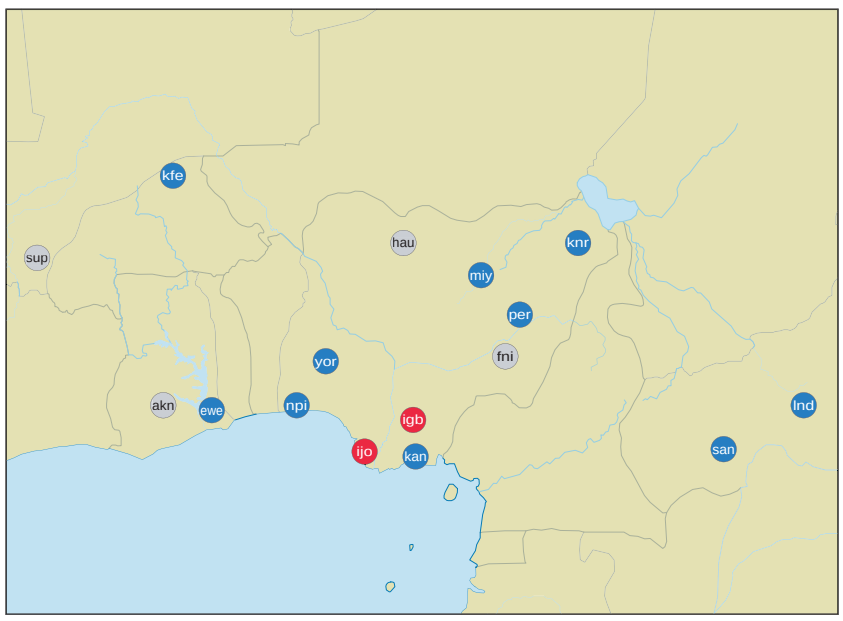
On the other hand, ‘when’ relations do not involve a number of semantic features that appear to be associated with the distribution of deranked verb forms across subordination relations cross-linguistically, such as semantic integration and sharing of participants between main and dependent event. ‘When’ relations do not involve semantic integration between the linked events. This disfavours the occurrence of deranked verb forms, because deranked verb forms **iconically** reflect semantic integration (Givón 1980 and 1990: ch. 13; see also Chapter 125). Also, ‘when’ relations do not involve sharing of participants between main and dependent event. As a result the identity of the participants of the dependent event cannot be recovered from the context. This disfavours the occurrence of verb forms not marked for person, as is often the case with deranked forms.

Thus, some functional principles motivate the occurrence of deranked ‘when’ clauses, while others motivate the occurrence of balanced ‘when’ clauses. This explains why both balanced and deranked ‘when’ clauses are widespread across the world’s languages.

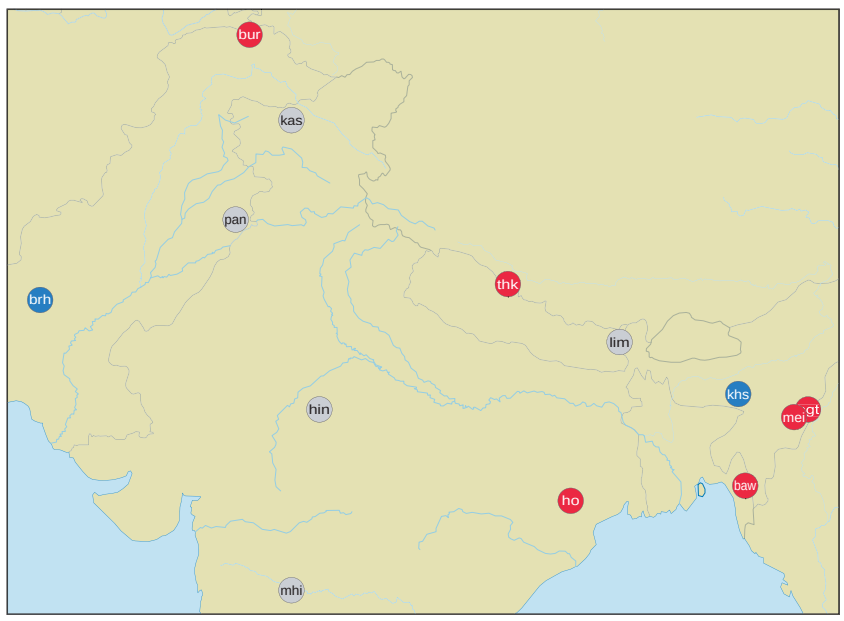
126 ‘When’ Clauses

SONIA CRISTOFARO

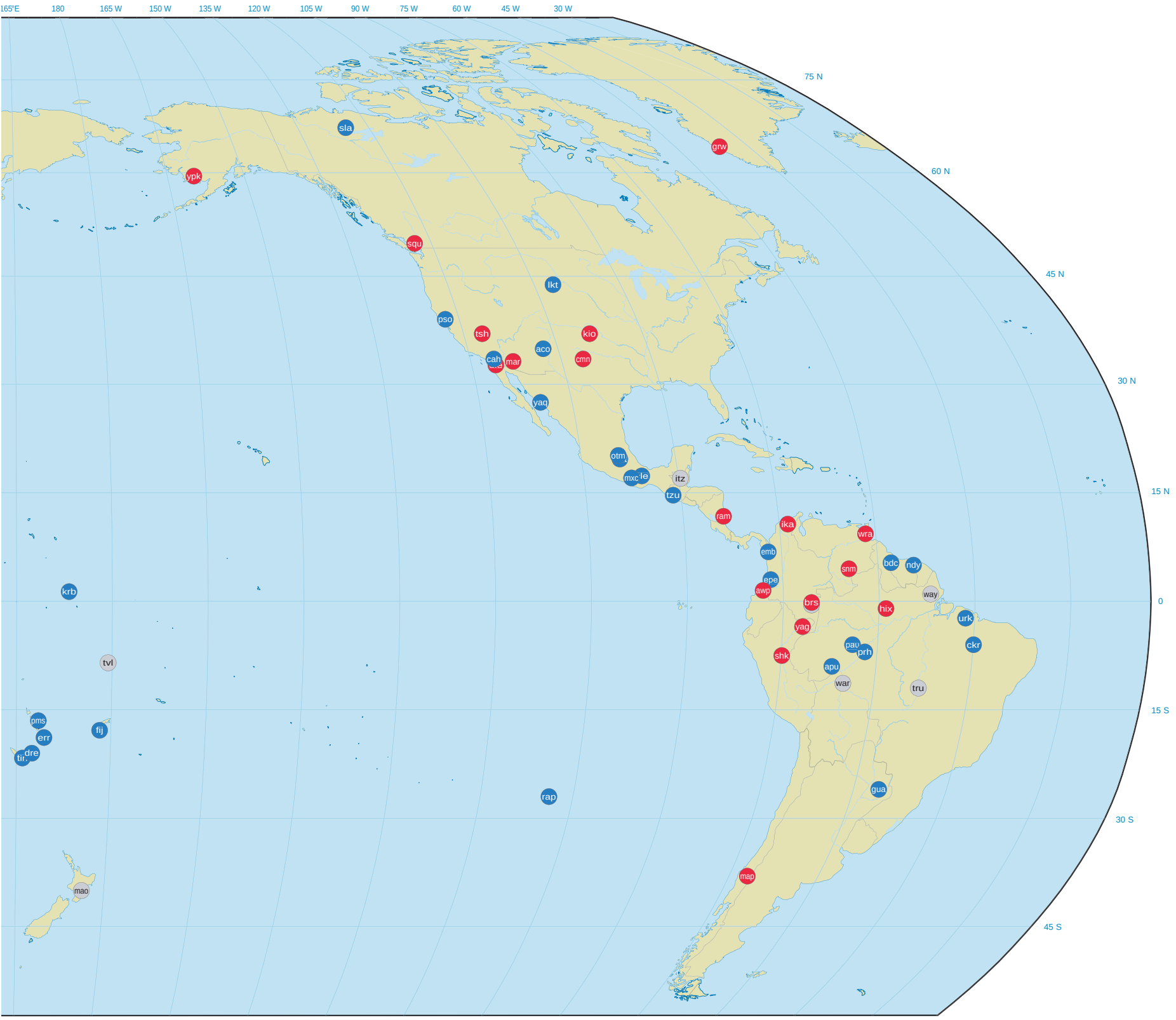


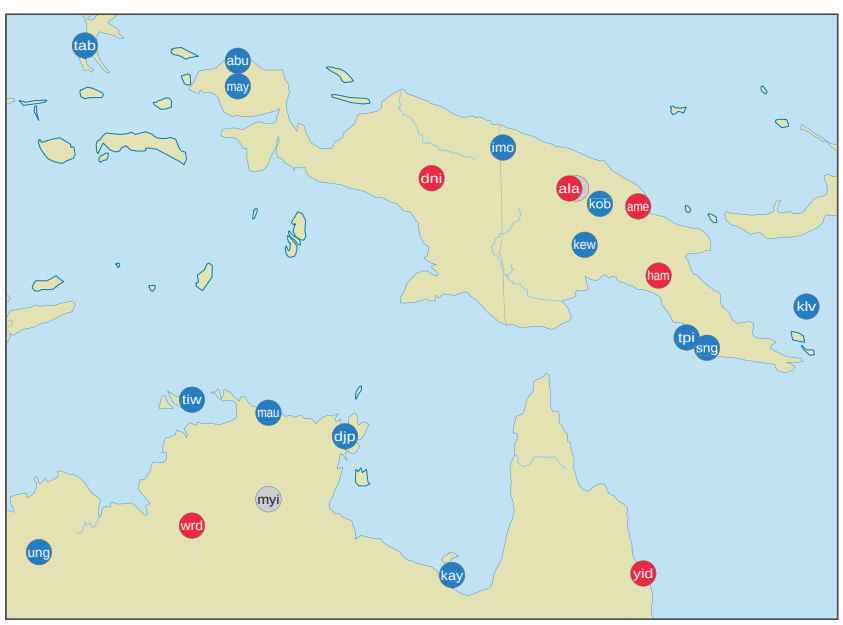


Nigeria and surrounding areas



northern India and surrounding areas





New Guinea and northern Australia

from southern Mexico to Columbia

127 Reason Clauses

SONIA CRISTOFARO

1 Defining the values

This chapter examines the form of the verb in reason clauses. Examples (1)–(3) provide instances of reason clauses in English (throughout the chapter, reason clauses will be enclosed in square brackets in the examples):

(1) *I couldn't go to Paris last week [because all the trains were booked out]*.

(2) *[Since it will be raining soon], it's better not to go out*.

(3) *[Being so busy], I couldn't do anything else*.

As was done for purpose clauses and 'when' clauses (Chapters 125 and 126), reason clauses are defined in functional, rather than morphosyntactic, terms. A reason construction is regarded as one encoding a causal relation between two events, such that one of the two (the event coded by the reason clause, or the dependent event) represents the reason for the other event (the main event) to take place.

In Cristofaro (2003: chs. 2, 6), the theoretical implications of the proposed functional definition of reason clauses are discussed at length, and some pragmatic criteria are provided to identify reason clauses under this definition.

The proposed functional definition of reason clauses encompasses both traditional cases of reason clauses (e.g. clauses introduced by specific reason conjunctions, as in (1)–(2), or nonfinite clauses, as in (3)) and clauses that might not count as such under traditional morphosyntactic criteria. One such case is provided by Canela-Krahô (Macro-Ge; Brazil):

(4) Canela-Krahô (Popjes and Popjes 1986: 139)
wa ha ma ajcahu, [i-mã hũpati]
1 FUT away run 1-TEMPRY 3.fear
'I will run away because I am afraid of it.'

In this sentence, the linked clauses are juxtaposed, and there is no grammatical marker signalling their linkage. The reason relation between the two events must be inferred on contextual grounds. However, the clause expressing the speaker's being afraid is regarded as a reason clause here, because it expresses the same kind of conceptual situation associated with clauses that are explicitly marked as expressing a reason relation between events in other languages.

Verb forms in reason clauses may be either balanced or deranked. The notions of **balancing** and **deranking** were originally introduced by Stassen (1985), and are defined in greater detail in Chapter 125. A balanced verb form is one that can occur in an independent declarative clause. A deranked verb form is one that cannot be used in independent declarative clauses. A deranked verb form may lack some or all of the categorial distinctions relevant to verbs in the language (such as tense, aspect, mood, or person agreement distinctions), or display special markers not used in independent clauses, e.g. special tense, aspect, mood, or person markers, nominalizers, case markers, or adpositions. The English examples in (1) and (2) provide instances of balanced reason clauses. Another example follows from Panjabi (Indo-European; India and Pakistan):

(5) Panjabi (Bhatia 1993: 75)
[cüüki ó bamaar ai], ó náïï aa
because he sick is he NEG come
sake-gaa
able-FUT.3M
'Because he is sick, he won't be able to come.'

In this sentence, the reason clause is introduced by the conjunction *cüüki* 'because', and the verb is fully inflected for the categories relevant to verbs in the language (tense, aspect, mood, and person).

Example (3) above provides an instance of a deranked reason clause: the verb is in the gerund form, and is not inflected for person and mood. Other examples of deranked reason clauses are provided in (6)–(8).

In example (6) from Kolyma Yukaghir (eastern Siberia), the verb in the reason clause is in the result nominal form. It is not inflected for tense, aspect, mood, and person, and bears the prolative case suffix *-gen*, as well as the result nominal affix *-o:l-*:

(6) Kolyma Yukaghir (Maslova 2003: 432)
*[tay marqil' eris' **ann'-o:l-de-gen]***
that girl badly speak-RNR-POSS-PROL
tabud-ek lem-mele
that-PRED eat-3SG
'Because that girl had spoken wrongly, he ate her.'

In example (7), from Italian, the verb in the reason clause is in the infinitive form, and is inflected for aspect but not for tense, mood, or person. In addition, it is introduced by the preposition *per* 'for':

(7) Italian
*[**Per** **av-ere** **detto** questo],*
for AUX-INF say.PTCP this
ha ricev-uto molte critiche.
AUX.PRES.IND.3SG get-PTCP many criticisms
'He was severely criticized for saying that.'

Example (8), from West Greenlandic (Eskimo; Greenland), illustrates the use of so-called dependent moods. These are verb forms that cannot occur in independent declarative clauses, but display all of the categorial distinctions relevant to verbs in the language. However, these distinctions are realized by means of special forms, different from those used in independent declarative clauses. In (8), the reason clause is in the so-called causative mood. The verb is inflected for tense, aspect, mood, and person, just like the forms used in independent declarative clauses. Tense and aspect are expressed as in independent declarative clauses, as witnessed by the use of the future affix *-ssa-* in both the main and the reason clause. However, mood and person are expressed by means of the causative affix *-mmat*, which cannot be used in independent declarative clauses.

(8) West Greenlandic (Fortescue 1984: 65)
[anurli-ssa-mmat] *aalla-ssa-nngil-agut*
be.windy-FUT-3SG.CAUS leave-FUT-not-1PL.IND
'Since it is going to be windy, we shan't leave.'

As was observed in connection with purpose and 'when' clauses (see Chapters 125 and 126 and references therein), it has been suggested that forms such as the causative mood in West Greenlandic might be regarded as balanced, because they display the same categorial distinctions as the forms used in independent declarative clauses. However, since they cannot themselves occur in independent declarative clauses, they will be regarded here as deranked.

In some cases, the verb forms used in reason clauses are identical to those used in independent clauses, but bear a clitic or an affixed conjunction. This is, for instance, the case in Trumai (isolate; Brazil), where the verb bears the affixed conjunction *-ak*:

(9) Trumai (Guirardello 1999: 404)
*iyi pumaţ-e [kawao-k iki-n-**ak**]*
3SG scream-3ABS wasp-ERG sting-3ABS-because
'He screamed because a wasp stung him.'

In such cases, one might want to regard the relevant verb form as deranked, because the complex verb + clitic/affixed conjunction could not occur in an independent declarative clause. In (9), however, the verb as such is structurally identical to a form used in

independent declarative clauses (in fact, the affix *–ak* is a suppletive form of the independent reason conjunction *iets* 'because': Guirardello 1999: 404). For this reason, it will be regarded here as balanced, and the same criterion will be applied to similar cases in other languages.

The balancing/deranking distinction overlaps with, but is not equivalent to, the distinction between finiteness and nonfiniteness. For discussion of the relevant issues, see Chapter 125.

For any given language, reason clauses can be coded by deranked verb forms only, by balanced verb forms only, or by either balanced or deranked verb forms. Map 127 shows the distribution of these three types in the world's languages.

1. Balanced	90
2. Balanced/deranked	37
3. Deranked	42
total	169

2 Geographical distribution

Languages with exclusively balanced reason clauses are overwhelmingly dominant in Central America, Africa, South-East Asia, and the Pacific islands.

In North America, most languages have exclusively deranked reason clauses. However, languages with exclusively balanced reason clauses and languages with both deranked and balanced reason clauses are also relatively frequent.

Languages with both deranked and balanced reason clauses are marginal in both Central and South America. In South America, languages with exclusively deranked reason clauses are approximately as frequent as languages with exclusively balanced reason clauses.

In Europe, most languages display both balanced and deranked reason clauses. The same situation is found in India. In Central Asia, languages with both deranked and balanced reason clauses are approximately as frequent as languages with exclusively balanced reason clauses. Languages with exclusively deranked reason clauses are found in the Caucasus, in Siberia, and in India.

In Australia, most languages display exclusively balanced reason clauses, but languages with exclusively deranked reason clauses are relatively common. On the other hand, languages with both deranked and balanced reason clauses are less frequent.

3 Theoretical issues

Most of the sample languages have exclusively balanced reason clauses. It should, however, be pointed out that languages with exclusively balanced reason clauses are concentrated in a number of geographical areas, such as Central America, Africa, mainland South-East Asia, Indonesia, and the Pacific islands. Most of these areas are the same ones that display a high concentration of languages with exclusively balanced purpose and 'when' clauses (Chapters 125 and 126). In other areas (such as North and South America, Europe, Central and South Asia, and Australia) balanced and deranked reason clauses are both widespread. This suggests that the distribution of languages with exclusively balanced reason clauses is areally and/or genetically biased, and both balanced and deranked reason clauses are widespread across the world's languages. It should also be observed, however, that the number of areas where balanced clauses are dominant is greater than in the case of other subordinate clause types, such as 'when' clauses (Chapter 126). This can be regarded as evidence that on the whole reason clauses display a greater tendency to be coded by means of balanced verb forms than 'when' clauses.

In fact, as is shown in Cristofaro (1998, 2003), an implicational hierarchy exists (the Subordination Deranking Hierarchy) such that if a language uses deranked verb forms for reason clauses, then it uses deranked verb forms for most of the other adverbial clause types, as well as for several types of complement clause and some types of relative clause (see also Chapter 125).

In Cristofaro (1998, 2003), a functionally motivated connection is established between some of the structural features of deranked verb forms and the semantic features of reason relations. Like purpose and 'when' relations, reason relations imply that the mood value of the dependent event is predetermined in that the dependent event is presented as factual. As a result, the mood value of the dependent event need not be overtly specified in the dependent clause, which leads to the use of verb forms not marked for mood, i.e. deranked forms. This reflects an **economic motivation**, the tendency not to express what is entailed or recoverable from the context (Haiman 1983, 1985; see also Chapters 125 and 126).

Also, reason relations allow the dependent event to be conceptualized as a thing rather than as a process (in Langacker's sense of these terms: see Langacker 1987a, 1987b and 1991, as well as Cristofaro 2003: chs. 6, 9). This may lead to the presence of nominal markers on the dependent verb (such as case markers, see for instance example (6) above), and nominal markers are a deranking device.

On the other hand, the occurrence of deranked forms is also related to semantic integration between events, in that deranked verb forms iconically reflect semantic integration (Givón 1980 and 1990: ch. 13; see also Chapter 125). However, this motivation does not hold for reason relations, because reason relations involve no semantic integration between events.

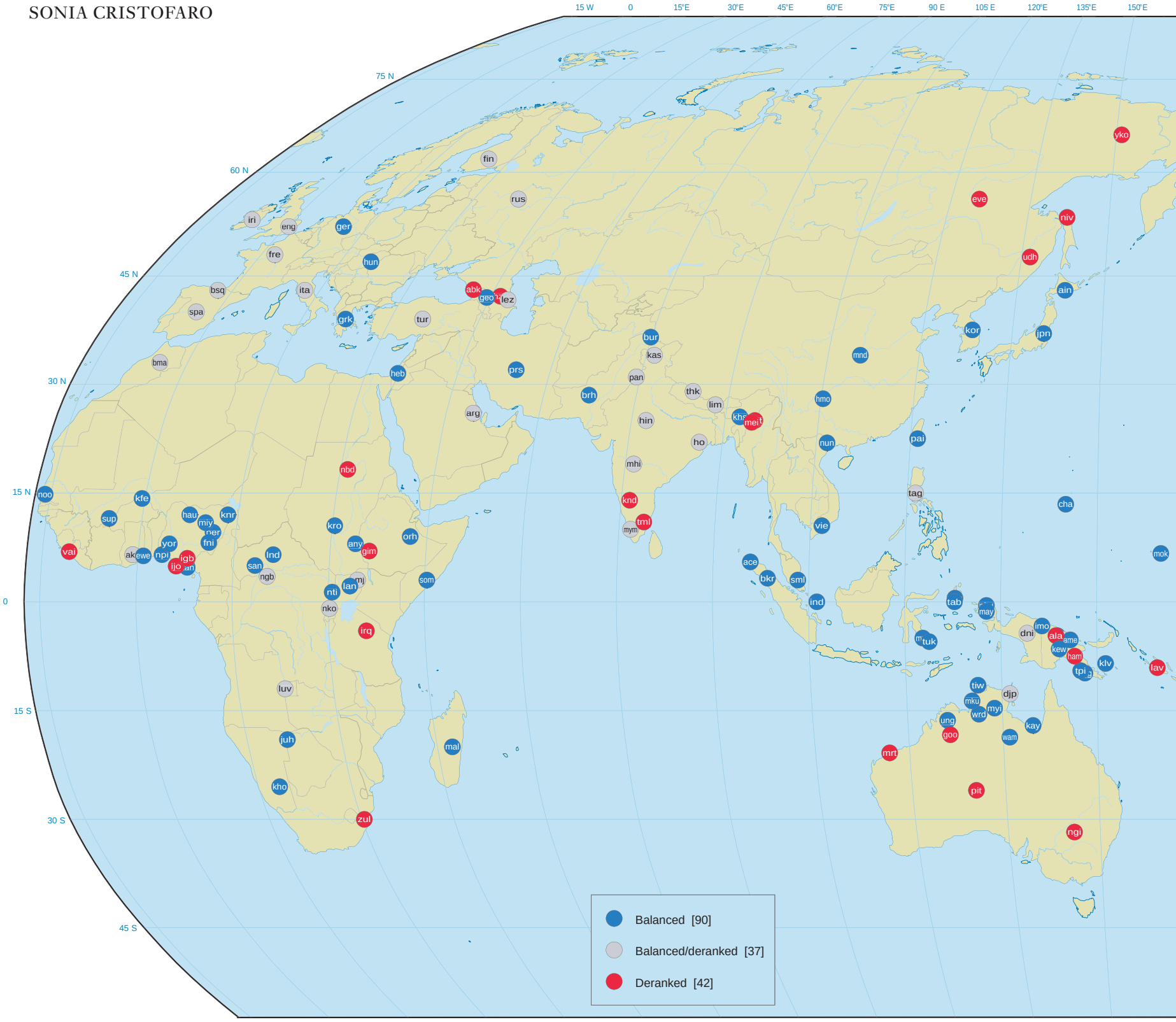
Also, reason relations do not involve predetermination of the time reference of the dependent event, nor sharing of participants between main and dependent event. This disfavours the occurrence of verb forms not specified for these categories, as is the case with deranked forms not marked for tense or person agreement.

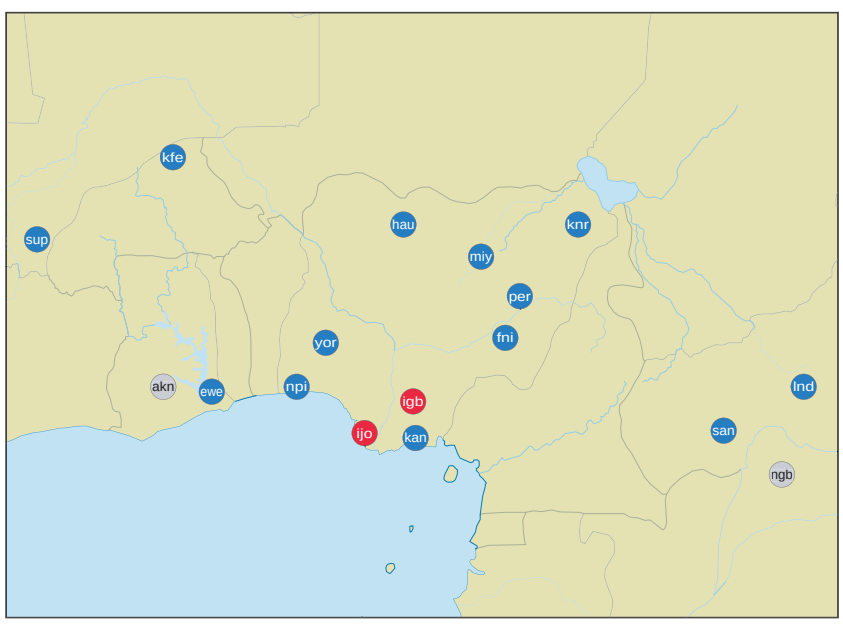
Finally, some deranked verb forms, such as subjunctives or conditionals, are devoted to the expression of nonfactuality (see the discussion in Chapters 125 and 126). These forms usually do not occur in reason clauses, as the dependent event is presented as factual.

The interaction of the various motivations just described may explain why both balanced and deranked reason clauses are widespread across the world's languages.

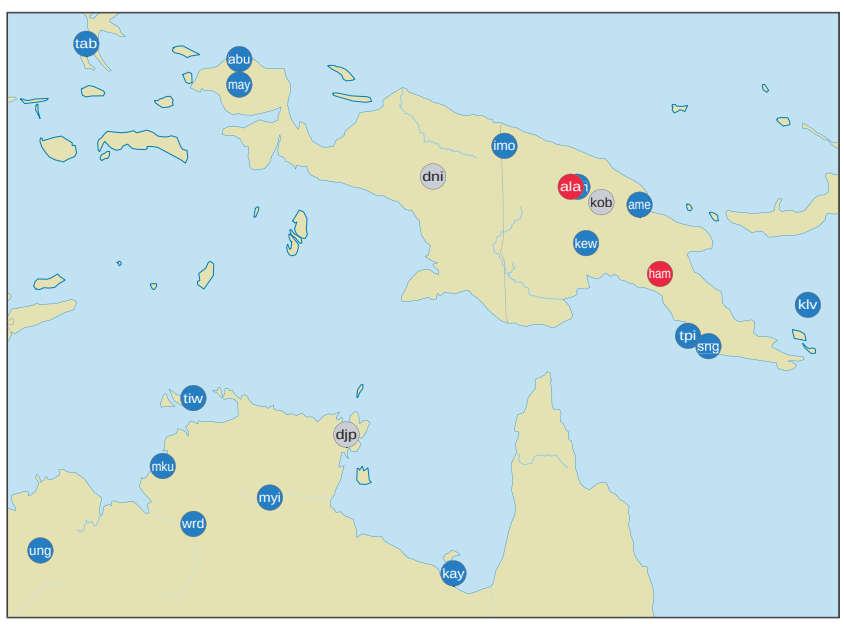
127 Reason Clauses

SONIA CRISTOFARO

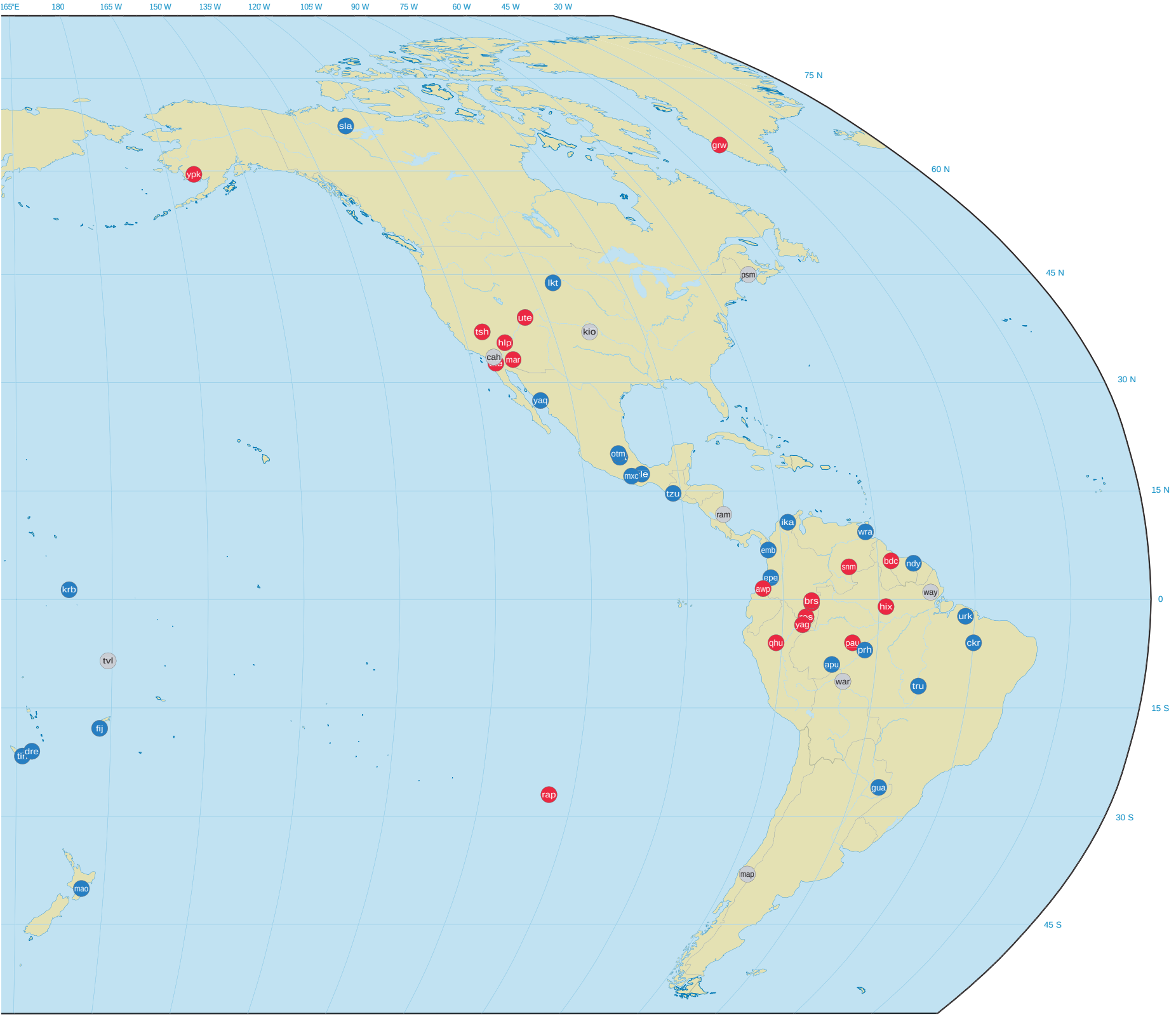


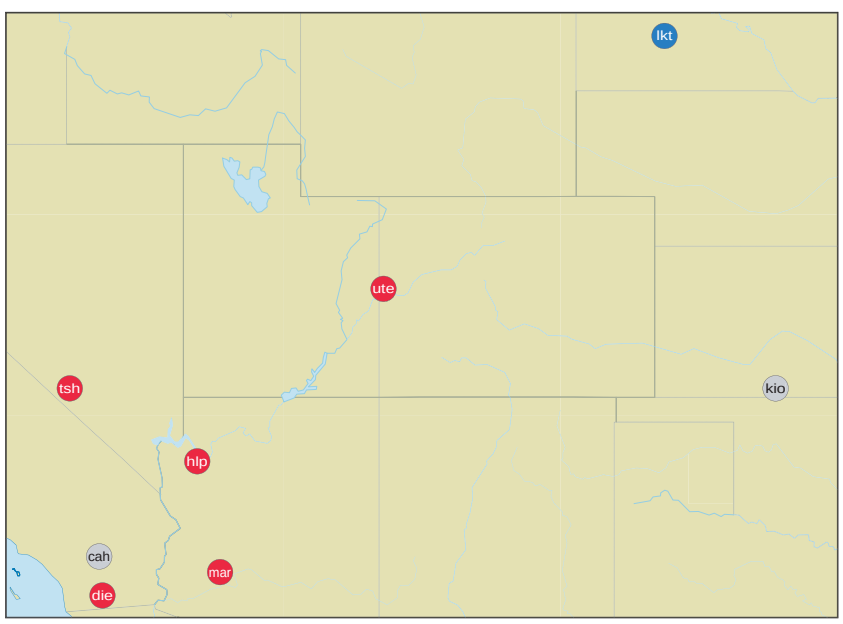


Nigeria and surrounding areas

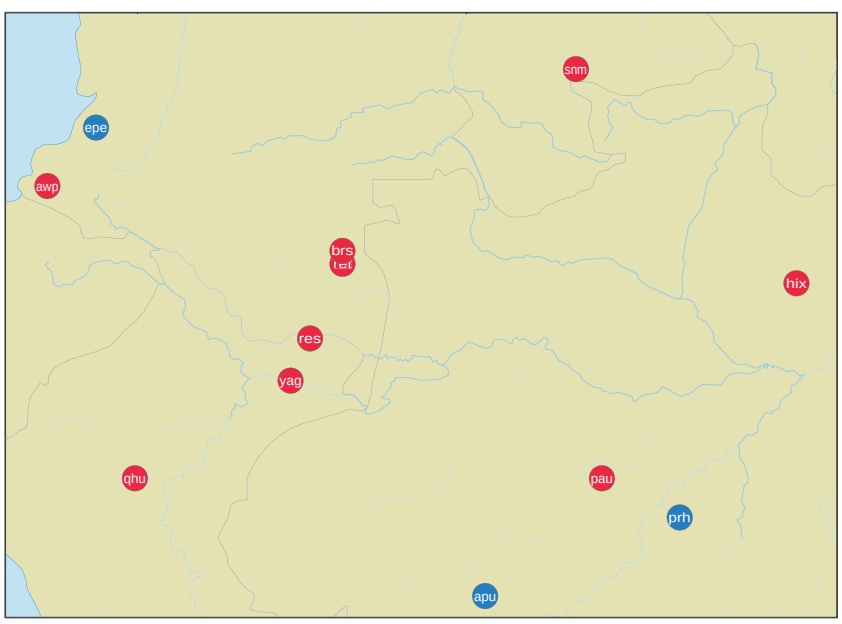


New Guinea and northern Australia





central-western United States



northwestern South America

128 Utterance Complement Clauses

SONIA CRISTOFARO

1 Defining the values

This chapter investigates the form of the verb in utterance complements, i.e. complement clauses introduced by utterance predicates such as 'say' or 'tell' (Noonan 1985: 110–13). Examples (1)–(2) provide instances of utterance complements in English (throughout the chapter, utterance complements will be enclosed in square brackets in the examples):

(1) *John said [that she came].*

(2) *John says [he will not do the job].*

Complement clauses are usually defined on the basis of syntactic criteria such as embedding, i.e. the fact that the complement clause functions as an argument of the main predicate (see for instance Noonan 1985: 46). However, as was done for purpose, 'when', and reason clauses (Chapters 125–7), utterance complements will be defined here in functional, rather than syntactic, terms. A complement construction is regarded as one expressing a particular relation between events, such that one event (the one coded by the main clause, or the main event) entails that another event (the one coded by the complement clause, or the dependent event) is referred to (for a detailed discussion of this definition, see Cristofaro 2003: chs. 2, 5). For instance, in (1) and (2), John's statement must refer to the occurrence of some event (coming, doing the job). Complement clauses may be introduced by different types of predicates (complement-taking predicates: cf. Noonan 1985), including utterance predicates.

In Cristofaro (2003: ch. 2), the theoretical implications of the proposed functional definition of complement clauses are discussed at length, and some pragmatic criteria are provided to identify complement clauses under this definition.

The proposed definition of complement clauses encompasses all of the traditional cases of utterance complements (embedded clauses, possibly introduced by a complementizer, as in (1) and (2)), plus some other cases that might not be regarded as such under traditional syntactic criteria. For instance, in a number of languages direct speech is the only means available to convey reported speech. This is the case in Kobon (Madang, Trans-New Guinea; Papua New Guinea):

(3) Kobon (Davies 1981: 1)
rol *Dusin* *laŋ* *nipe* *ip*
yesterday Dusin above 3SG ACC.1SG
hag-öp / *hag-öm* *[yad* *ram*
say-PERF.3SG say-SS.3SG 1SG house
ar-ab-in *a* *g-öp]*
go-PRES-1SG QUOT do-PERF.3SG
'Yesterday at Dusin he said to me, "I am going home." ' /
'Yesterday at Dusin he told me that he was going home.'

Other languages display constructions that are not specifically devoted to the expression of either direct or indirect speech, but can be used for both, with a shift in the reference of deictics. This can be seen from the English translations of the following Pirahã (Mura; Brazil) example:

(4) Pirahã (Everett 1986: 269)
hi *gái-sai* *[xahóápáti* *ti* *xi* *aagá-hóág-a]*
3 say-NMLZ Xahoapati 1 hunger have-INGR-REMOTE
'Xahhoapati said, "I am hungry." '/
'Xahhoapati said that I am hungry.'

A direct report differs from an indirect report in a number of structural and functional ways, and is not an argument of the utterance verb introducing it (Munro 1982). However, if direct speech (or a construction making no difference between direct and indirect speech, such as the one in (4)) is the only means available in a language to convey reported speech, it should be assumed that it may also express the conceptual situation expressed by indirect speech in other languages. If this were not the case, the language would have no means to express the relevant conceptual situation (see on this point Cristofaro 2003: chs. 2, 5). Therefore, direct-speech clauses are regarded as utterance complements here, unless the language also has some indirect speech construction. This is possible because the proposed functional definition of utterance complements is independent of any particular structural feature of the relevant clauses, and only refers to the conceptual relation between the linked events.

Verb forms in utterance complements may be either balanced or deranked, and the corresponding clauses will henceforth be called balanced utterance complements and deranked utterance complements. The notions of **balancing** and **deranking** were originally introduced by Stassen (1985), and are defined in greater detail in Chapter 125. A balanced verb form is one that can occur in an independent declarative clause, e.g. an indicative verb form. The utterance complements in examples (1)–(4) above are balanced. The verb is inflected for the same categories as the forms used in independent declarative clauses (for instance, tense, aspect, and mood in (1), (2), and (4), and tense, aspect, mood, and person in (3)), and these categories are expressed in the same way as in independent declarative clauses.

A deranked verb form is one that cannot be used in independent declarative clauses. A deranked verb form may lack some or all of the categorial distinctions relevant to verbs in the language (such as tense, aspect, mood, or person agreement distinctions), or display special markers not used in independent declarative clauses, e.g. special tense, aspect, mood, or person markers, nominalizers, case markers, or adpositions. Example (5) from Maricopa (Yuman; Arizona) provides an instance of a deranked utterance complement. The verb is not inflected for tense, aspect, and mood, and bears the locative case marker *-k*:

(5) Maricopa (Gordon 1986: 130)
nyaa *'i-m* *[Pam-sh* *mdiily* ***maa-k]*** *'i-sh*
I say-REAL Pam-SUBJ bread eat-LOC say-PERF
'I said that Pam ate the bread.'

In example (6), from Italian, the verb is in the infinitive, and is not inflected for tense, mood, or person. In addition, it is introduced by the preposition *di* 'of'. Infinitival utterance complements are used in Italian when main and dependent clauses share their subject.

(6) Italian
Dic-e *di* *non* *pot-er* *venire*.
say-PRES.IND.3SG of NEG can-PRES.INF come-PRES.INF
'S/he says s/he cannot come.'

Another case of deranked utterance complements is represented by the use of so-called dependent moods. As was pointed out in Chapters 125–7, dependent moods are verb forms displaying the same categorial distinctions as the verb forms used in independent declarative clauses. However, these distinctions are realized by means of special forms not used in independent declarative clauses; thus dependent moods cannot occur in independent declarative clauses. For instance, in the following West Greenlandic (Eskimo; Greenland) example, the verb is in the so-called contemporative mood. Tense and aspect are expressed as in independent clauses, but mood and person are expressed by means of special affixes that cannot be used in independent declarative clauses:

(7) West Greenlandic (Fortescue 1984: 40)
uqar-sinnaa-vunga *tama-asa* ***[uuqattaar-sima-llugit]***
say-can-1SG.IND all-3PL try-PERF-1SG.3PL.CONTEMP
'I can say that I have tried them all.'

As Stassen (1985: 338–9) observes, if a verb form is inflected for the same categories as the forms used in independent declarative clauses, one might want to regard it as balanced. However, since dependent moods cannot themselves be used in independent declarative clauses, they will be regarded as deranked here.

As was pointed out in Chapters 125–7, the distinction between balancing and deranking is not equivalent to the distinction between finiteness and nonfiniteness, although the two overlap to a significant extent. For discussion of the relevant issues, see Chapter 125.

For any given language, utterance complements can be coded by balanced verb forms only, by deranked verb forms only, or by either balanced or deranked verb forms. Map 128 shows the distribution of these three types in the world's languages.

1. Balanced	114
2. Balanced/deranked	18
3. Deranked	11
total	143

2 Geographical distribution

Languages with exclusively balanced utterance complements are overwhelmingly dominant in the language sample as a whole. They are also dominant in all geographical areas, except for an area extending from Europe in the west to Siberia in the east. Although several languages with exclusively balanced utterance complements are found in Europe, this area displays a high number of languages with both deranked and balanced utterance complements, as well as a pocket of languages in Siberia with exclusively deranked utterance complements.

Apart from this area, languages with both deranked and balanced utterance complements are also found in North America, Africa, and India, but they are marginal in all of these areas. Languages with exclusively deranked utterance complements can be found in North and South America, India, and Australia. They are marginal in all of these areas. However, South America displays a higher concentration of languages with exclusively deranked utterance complements than the other just-mentioned areas.

3 Theoretical issues

The most significant fact about the distribution of balanced and deranked utterance complements across the language sample is the overwhelming dominance of languages with exclusively balanced utterance complements. These languages are widespread, and in most cases dominant in all genetic families and geographical areas (with the exception of Siberia). This suggests that deranked utterance complements are disfavoured cross-linguistically.

In Cristofaro (1998, 2003) it is argued that the occurrence of deranked verb forms is related to a number of semantic features of particular subordination relation types, and utterance complements do not display these features. For instance, deranked verb forms are often not marked for tense, aspect, mood, or person agreement. As a result, they tend to be used when the time reference, the aspect and mood value, and the participants of the dependent event are predetermined, and thus need not be overtly specified. However, utterance complements do not involve predetermination of the time reference or the participants of the dependent event (for instance, examples (1) and (2) show that utterance complements may have different time reference, and may or may not share a participant with the main clause; on this point see also Noonan 1985). This disfavours the occurrence of verb forms not marked for tense or person agreement. Also, the occurrence of deranked verb forms is related to semantic integration between the linked events (in the sense defined in Givón 1980, 1990: ch. 13; see the discussion in Chapter 125). However, utterance complements involve no semantic integration between the linked events – in fact, the subordination relation here is not established between two events, but between an event (the main event) and a propositional content referring to another event (see on this point Dik 1989, 1997a, and 1997b; Hengeveld 1989 and 1990; Siewierska 1991).

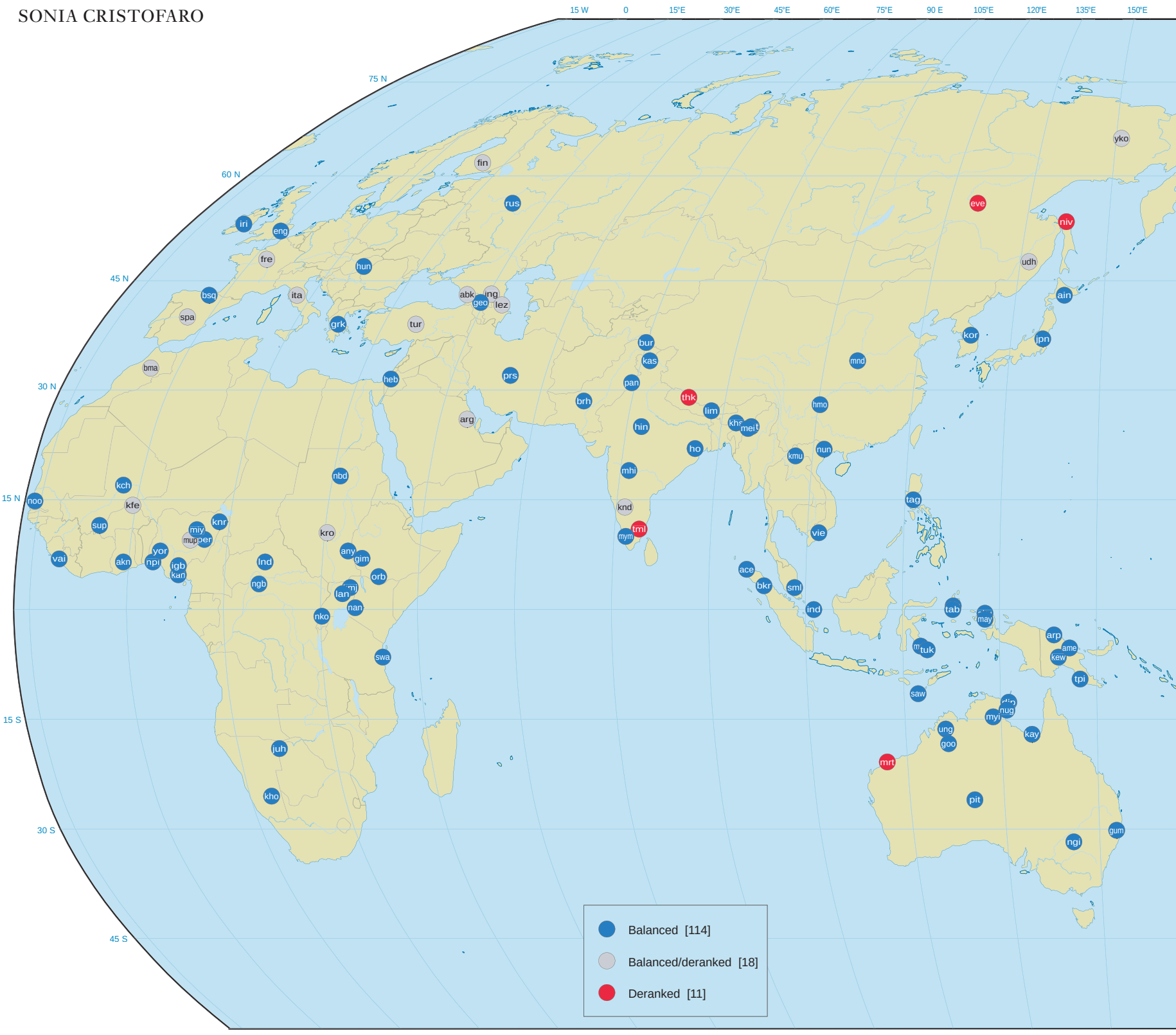
The occurrence of deranked verb forms is also related to the ability of the dependent event to be conceptualized as a thing rather than as a process (in Langacker's sense of these terms: see Langacker 1987a, 1987b, and 1991, as well as Cristofaro 2003: chs. 6, 9). Conceptualization of the dependent event as a thing rather than as a process leads to the presence of nominal properties on the verb, such as case markers or adpositions, as well as the absence of typical verbal properties such as tense, aspect, and mood distinctions. All of these phenomena are characteristic of deranking. However, utterance complements do not seem to easily allow conceptualization of the dependent event as a thing, and this disfavours the occurrence of verb forms displaying nominal properties, or lacking verbal properties.

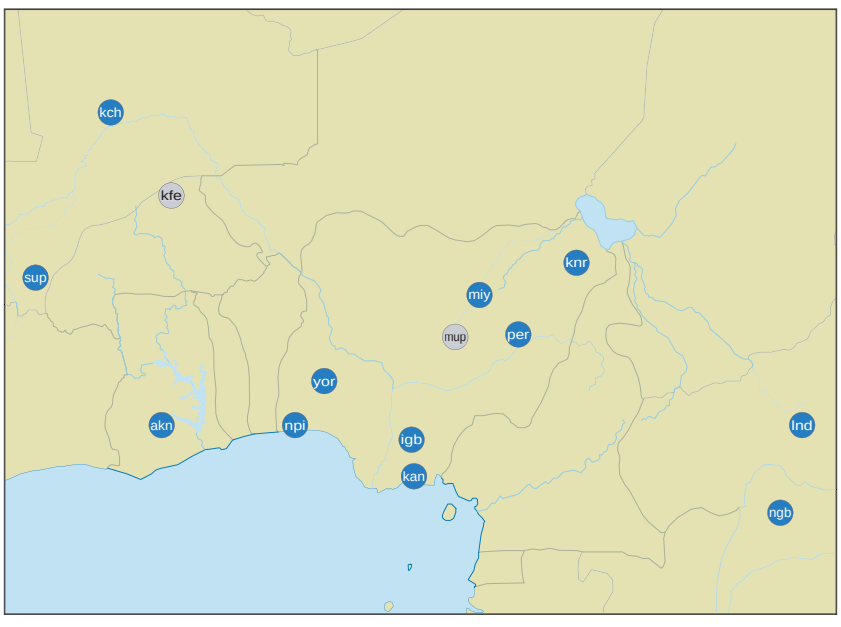
Finally, the occurrence of some deranked forms such as subjunctives and the like is related to the dependent event being presented as unrealized. However, utterance complements involve no implication about whether or not the dependent event is realized. This disfavours the occurrence of verb forms devoted to the expression of irrealis, such as subjunctives.

All this may explain why deranked utterance complements are disfavoured cross-linguistically. In fact, as is shown in Cristofaro (1998, 2003), an implicational hierarchy (the Subordination Deranking Hierarchy) exists such that if a language uses deranked verb forms for utterance complements then it will use deranked verb forms for all of the other complement types, as well as for most types of relative and adverbial clauses.

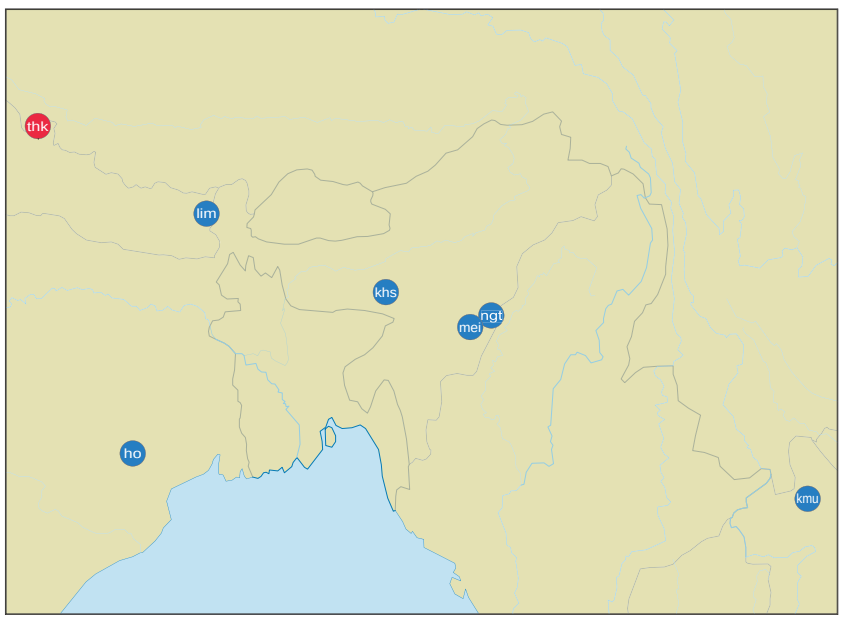
128 Utterance Complement Clauses

SONIA CRISTOFARO

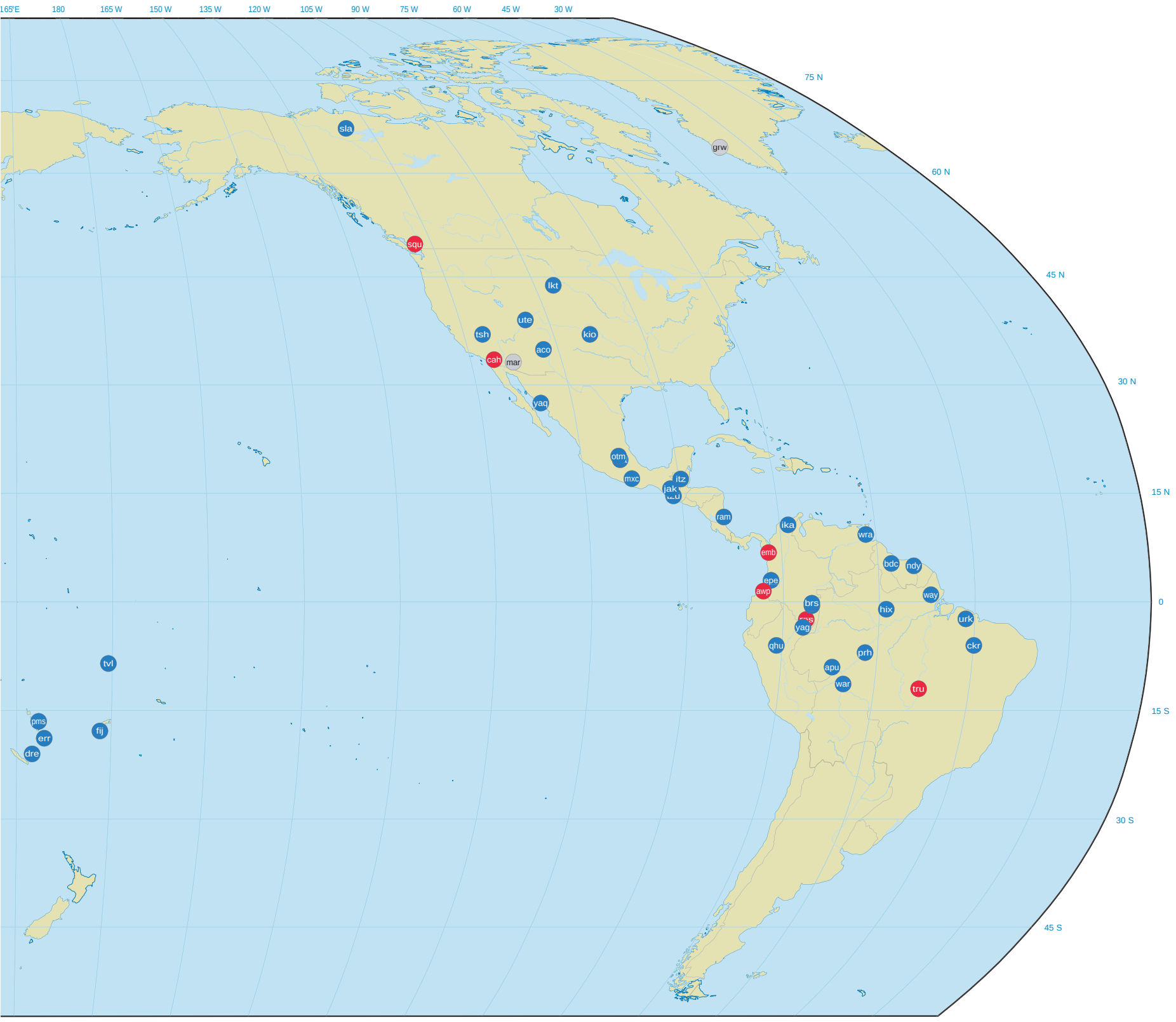


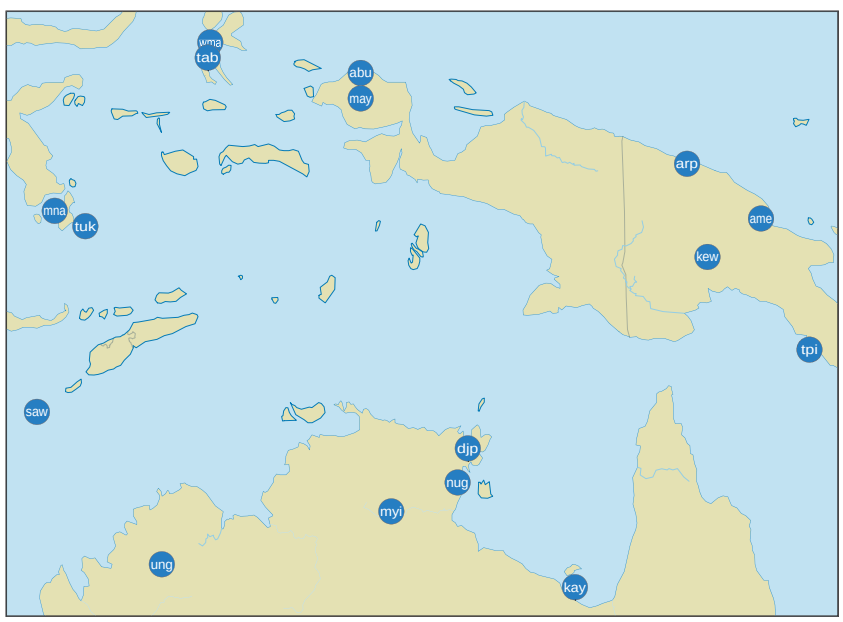


Nigeria and surrounding areas

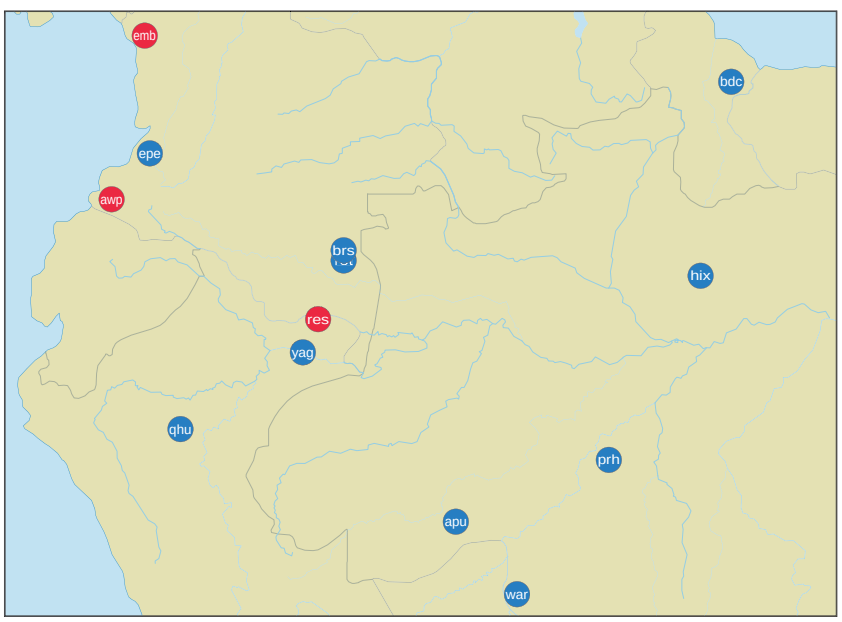


eastern India and surrounding areas





New Guinea and northern Australia



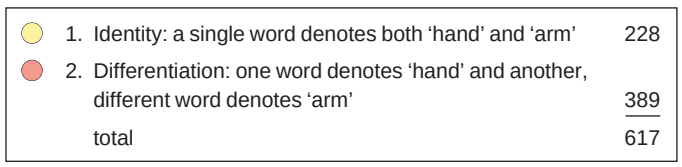
northwestern South America

129 Hand and Arm

CECIL H. BROWN

1 Defining the values

This map shows the distribution of the two primary ways in which languages lexically recognize major segments of the human upper limb. The major segments are 'hand', from the fingertips to the wrist, and 'arm', from the fingertips or from the wrist to the shoulder. Two values are represented:



English is an example of a **differentiating** (type two) language, with *hand* and *arm*. Another example is Ngawun (Pama-Nyungan; Queensland, Australia), with *marl* 'hand' and *palkal* 'arm'.

If there is no 'arm' term in a language, but one or more words denote respectively one or more major segments of 'arm' other than 'hand' such as the forearm and the upper arm, the language is also included in type two. An example is Chai (Nilo-Saharan; Ethiopia), having a word for 'hand', *síyó*, and a word for both 'forearm', *múní*, and 'upper arm', *yíró*, while lacking a term for 'arm'. Differentiating languages such as Chai that lack a word for 'arm' are rare among languages sampled for the map.

Examples of **identity** (type one) come from Lonwolwol (Oceanic; Vanuatu), Czech, and Gurma (Gur, Niger-Congo; Burkina Faso), respectively showing the terms *va:*, *ruka*, and *nu*, all of which denote both 'hand' and 'arm'.

Some languages included on the map are type one but nonetheless also have a term for 'hand' and/or one for 'arm' that does not show 'hand/arm' polysemy. For example, Bambara (Mande; Mali) uses *bolo* for both 'hand' and 'arm', but has an alternate term, *tègè,* used for 'hand' (and also denoting 'palm' and 'foot'). Semai (Mon-Khmer; Malay Peninsula) shows *tek*, which polysemously denotes both 'hand' and 'arm', and also *kengrit*, which designates only 'arm'. Jicarilla Apache (Athapaskan; New Mexico) denotes both 'hand' and 'arm' with *gan* but also designates the former with *l-lá* and the latter with *ganí*. The convention followed here is that if a language shows the same term for both 'hand' and 'arm,' type one (identity) is judged present, even if there are other terms in its lexicon denoting just 'hand' or 'arm'.

In some languages, a word denoting 'hand' also extends polysemously to a larger upper limb segment other than 'arm' that includes 'hand' as a part. For example, the Indonesian word for 'hand', *tangan*, also denotes the forearm (with *lengan* referring to 'arm'). Indonesian is judged to be a differentiating language. In Kadazan (Austronesian; Borneo), *hongon* denotes both 'hand' and 'forearm', but since it also refers to 'arm', the language is judged to be type one.

2 Geographical distribution

In the language sample used for this map, type two (differentiating) languages are about one and a half times more common than type one languages. There is a geographical distributional pattern: the percentage of type one languages occurring at latitudes closer to the equator is considerably greater than the percentage of type one languages found at latitudes more distant from the equator. In addition, among both type one and type two languages considerably removed from the equator (north of 35°59′N and south of 9°59′S), the percentage of type two languages is substantially greater than the percentage of type one languages. In other words, type one languages tend to occur more frequently nearer the equator and less frequently away from the equator. The statistical association between latitudinal location and values is strongly positive (gamma = .61, p < .001, see Table 1).

Table 1 Association between latitude and 'hand'/'arm' values

Latitude	Values	
	Differentiation	**Identity**
70°N–36°N, 10°S–55°S	231	60
35°59′N–9°59′S	158	168
gamma = .61	p < .001	N = 617

A geographical summary is as follows. Of the map's sixty languages native to Australia (and Tasmania), all but one, Yulparija, are differentiating. Among native languages of South America, differentiation is the dominant pattern. Differentiation is also dominant in Europe. Native languages of most of North America north of Mexico are differentiating with the exception of those of the US South-West and southern California, which mainly show identity. Languages manifesting identity dominate in southern Mexico, Central America, and the equatorial zones of Africa, Asia, and the Pacific.

3 Genealogical distribution

In some instances, genealogically related languages distributed broadly across latitudes show heterogeneous values (both type two and type one). In such cases, often languages closer to the equator tend more strongly to show identity than their closely related sister languages located further from the equator (which tend to show differentiation). For example, this is seen in the distribution of values across Uto-Aztecan languages and dialects (North America) ranked by latitude from north to south (cases of identity are given in bold):

Table 2 Uto-Aztecan languages ranked by latitude

41°N	Shoshone	2
39°N	Northern Paiute	2
38°N	Comanche	2
37°N	Tümpisa Shoshone	2
37°N	Ute	2
36°N	**Hopi**	**1**
36°N	Kawaiisu	2
35°N	Chemehuevi	2
33°40′N	**Cahuilla**	**1**
33°30′N	**Luiseño**	**1**
32°N	**O'odham**	**1**
29°N	**Tarahumara (Central)**	**1**
28°10′N	Mayo	2
28°N	Tubar	2
22°30′N	**Cora**	**1**
21°20′N	**Huichol**	**1**
19°30′N	**Sierra de Zacapoaxtla**	**1**
19°N	**Tetelcingo Nahuatl**	**1**
18°30′N	**Pajapan Nahuatl**	**1**
18°N	Huauchinango Nahuatl	2
18°N	Xalitla Nahuatl	2

Only 12.5 per cent of Uto-Aztecan languages spoken north of 34°N show identity, while 69.2 per cent of these languages located south of 34°N show identity. Conversely, 87.5 per cent of Uto-Aztecan languages north of 34°N show differentiation, while only 30.8 per cent south of 34°N show differentiation.

A similar pattern is observed for Benue-Congo languages of Africa:

Table 3 Benue-Congo languages ranked by latitude

6°40′N	**Igbo**	**1**
6°10′N	Babungo	2
5°N	**Efik**	**1**
4°25′N	**Ogbronuagum**	**1**
3°N	**Yaka**	**1**
0°30′S	Nkore-Kiga	2
4°S	**Kituba**	**1**
5°S	**Nyamwezi**	**1**
5°S	**Kongo**	**1**
6°30′S	**Swahili**	**1**
10°S	Bemba	2
18°S	**Shona**	**1**
20°S	Kalanga	2
22°S	Tswana	2
30°S	Zulu	2

And the pattern holds for Athapaskan languages of North America:

Table 4 Athapaskan languages ranked by latitude

67°N	Slave	2
67°N	Kutchin	2
65°N	Koyukon	2
65°N	Lower Tanana	2
63°N	Northern Tutchone	2
63°N	**Tanaina**	**1**
62°N	Ahtna	2
62°N	Degexit'an	2
60°N	Slavey	2
59°N	Chipewyan	2
53°20′N	Carrier	2
41°20′N	Hupa	2
38°N	Navajo	2
36°N	**Jicarilla Apache**	**1**
34°N	**Western Apache**	**1**
32°N	Chiricahua Apache	2

And for Indo-European languages of South Asia:

Table 5 Some Indo-European languages ranked by latitude

34°N	Kashmiri	2
31°N	Panjabi	2
26°N	**Maithili**	**1**
21°N	**Kotia Oriya**	**1**
19°N	**Marathi**	**1**
7°N	**Sinhala**	**1**
6°N	**Vedda**	**1**

Genealogically related languages broadly distributed from north to south do not always show heterogeneous values. For example, languages of the Australian phylum, with one exception, are uniformly differentiating despite distribution across roughly thirty degrees of latitudinal space. Another example involves Tupian languages of South America (Chiriguano, Guaraní, Tapieté, Mundurukú, Omagua, Guajajara, Tupi, Urubú-Kaapor, and Wayampi), which are distributed across twenty-five degrees of latitude, but are found to be uniformly differentiating.

Other groups of genealogically related languages show homogeneous values. All seven Numic languages (Shoshone, Northern Paiute, Comanche, Tümpisa Shoshone, Ute, Kawaiisu, and Chemehuevi) among the Uto-Aztecan languages listed in Table 2 are differentiating; all of these are spoken in the same region (Great Basin and abutting areas). The twelve Arawakan languages of South America (Amuesha, Arawak, Axininca Campa, Baniwa, Bare, Campa Pajonal Asheninca, Machiguenga, Nomatsiguenga, Piapoco, Piro, Resigaro, and Goajiro), spoken close to the equator between 13°N and 13°S, are all, with the single exception of Nomatsiguenga, differentiating.

Other genealogical groups show heterogeneous values but, unlike the Uto-Aztecan, Benue-Congo, Athapaskan, and Indo-European groups cited above, show no special distributional patterns. Of the twenty-one Altaic languages of the sample, which range from 57°N to 34°20′N, fifteen show identity (Tatar, Khakas, Kazakh, Solon, Udihe, Khalkha, Noghay, Kalmyk, Kumyk, Manchu, Azerbaijani, Uzbek, Ainu, Turkmen, and Mangghuer) and six are differentiating (Evenki, Buriat, Karachay-Balkar, Kirghiz, Turkish, and Moghol). The eighteen Tucanoan languages of South America, all spoken within one degree of latitude of the equator, are heterogeneous. Thirteen of these are differentiating (Barasano, Carapana, Desano, Guanano, Cubeo, Macuna, Northern Barasano, Secoya, Siriano, Retuarã, Tatuyo, Tuyuca, and Yuruti), and five (Koreguaje, Orejón, Siona, Piratapuyo, and Tucano) show identity.

Most of the Indo-European languages of the map, most of which are spoken at latitudes considerably removed from the equator (in areas north of 31°N), show differentiation. (Indo-European languages spoken south of 31°N show identity; see above discussion.) The Slavic and Baltic languages of Indo-European are exceptional. All of these (Belorussian, Czech, Lithuanian, Polish, Russian, Serbian-Croatian, Slovak, and Ukrainian) are spoken in regions north of 43°N, yet all but one (Polish) show identity.

4 Explanatory framework

The association between values and latitudinal location was first observed by Witkowski and Brown (1985). These authors propose that the existence of extensive wearing apparel in human groups negatively influences the occurrence of upper limb polysemy. The presence of tailored clothing covering the arms greatly increases the distinctiveness of arm parts and renders more likely their labelling by separate terms. In addition, ancillary apparel such as gloves and mittens also increases the salience of arm parts. Since nonequatorial zones where cold weather is frequent are usually associated with the presence of tailored clothing and other arm gear, languages spoken in these areas are significantly more inclined to lexically distinguish 'hand' and 'arm' than those spoken in equatorial zones.

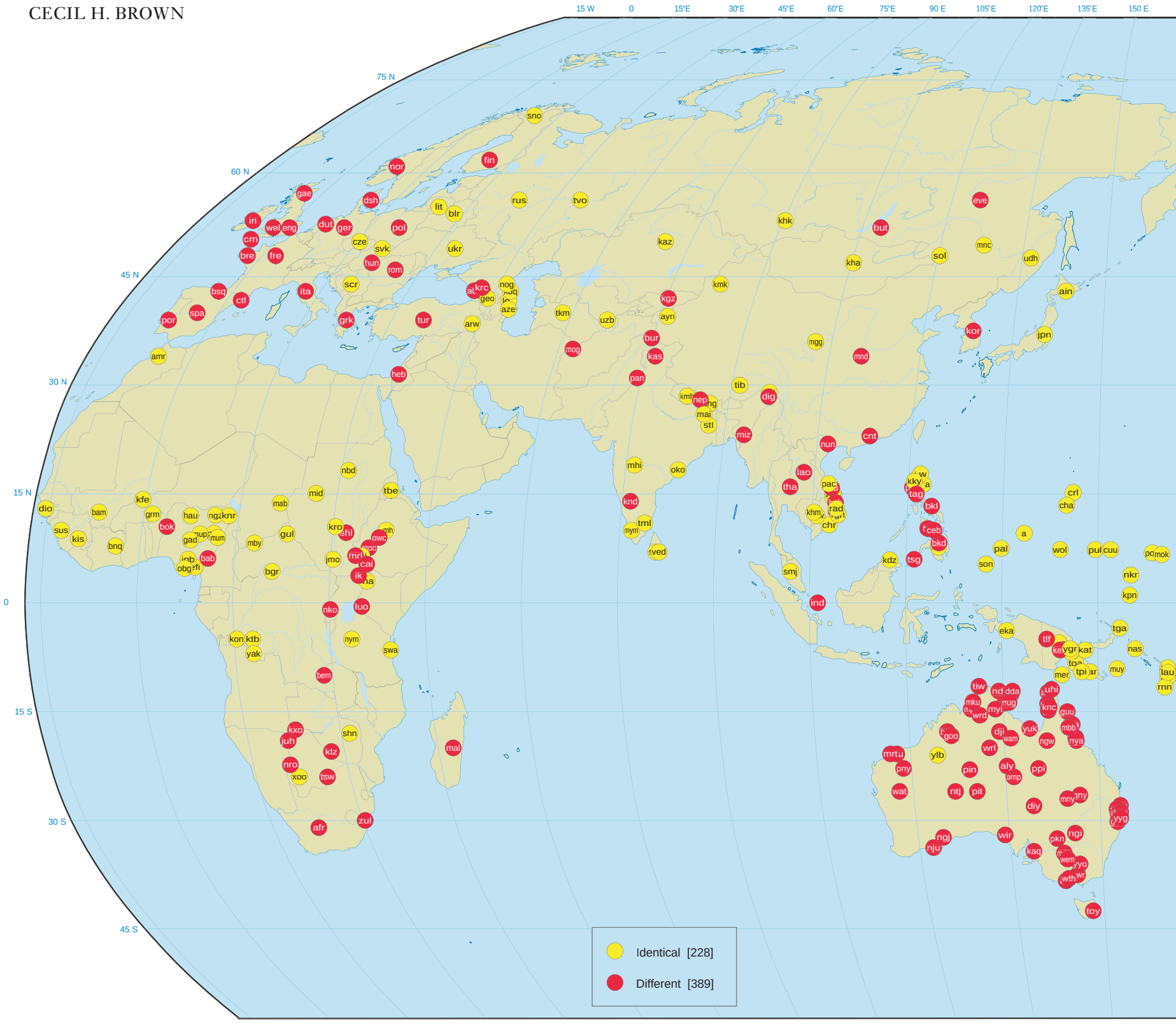
Latitudinal location is only an indirect index of temperature. While there is a strong correlation between values and latitudinal location (see Table 1), there may be an even stronger association between values and average annual temperature if the Witkowski-Brown hypothesis is correct. Unfortunately, data of a detailed nature facilitating plotting of values (type one and type two) across isotherms is not available at this stage in the development of world climatological science. Another potentially fruitful investigatory strategy would be to cross-tabulate values against the tailoring technologies of peoples who speak each of the 617 languages of the sample—an enormous research effort this author must leave to future investigators.

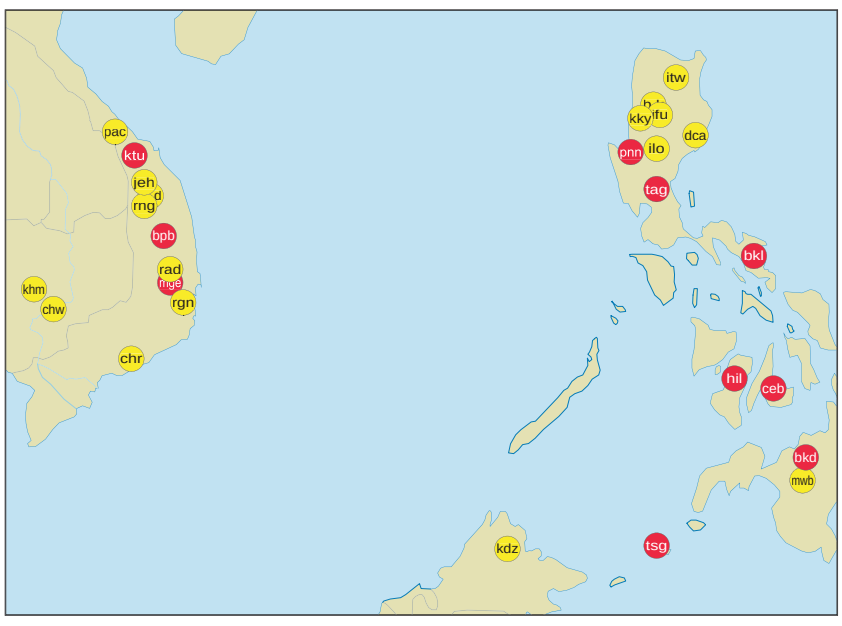
5 'Finger', 'hand', and 'arm'

Map 130 plots the distribution of languages using a single term to denote both 'finger' and 'hand'. Some languages sampled on the current map are three-ways polysemous, using a single term to denote all three of the referents, 'finger', 'hand', and 'arm'. These are Cahuilla (Uto-Aztecan; California), Cocopa (Yuman; Arizona), Comecrudo (isolate; Mexico), Mesa Grande Diegueño (Yuman; California), Eudeve (Uto-Aztecan; Mexico), Hawaiian (Polynesian; Hawaii), Pacoh (Mon-Khmer; Vietnam and Laos), Quileute (Chimakuan; Washington State), Seri (isolate; Mexico), Tahitian (Polynesian), Tuamotuan (Polynesian; French Polynesia), Western Apache (Athapaskan; Arizona), Yagaria (Trans-New Guinea; Papua New Guinea), and Yukpa (Carib; Colombia and Venezuela).

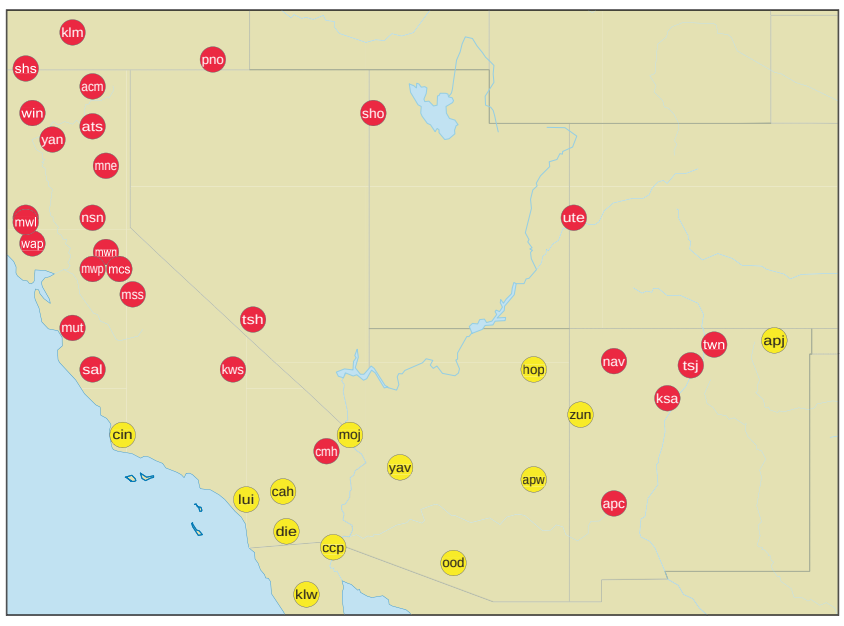
129 Hand and Arm

CECIL H. BROWN

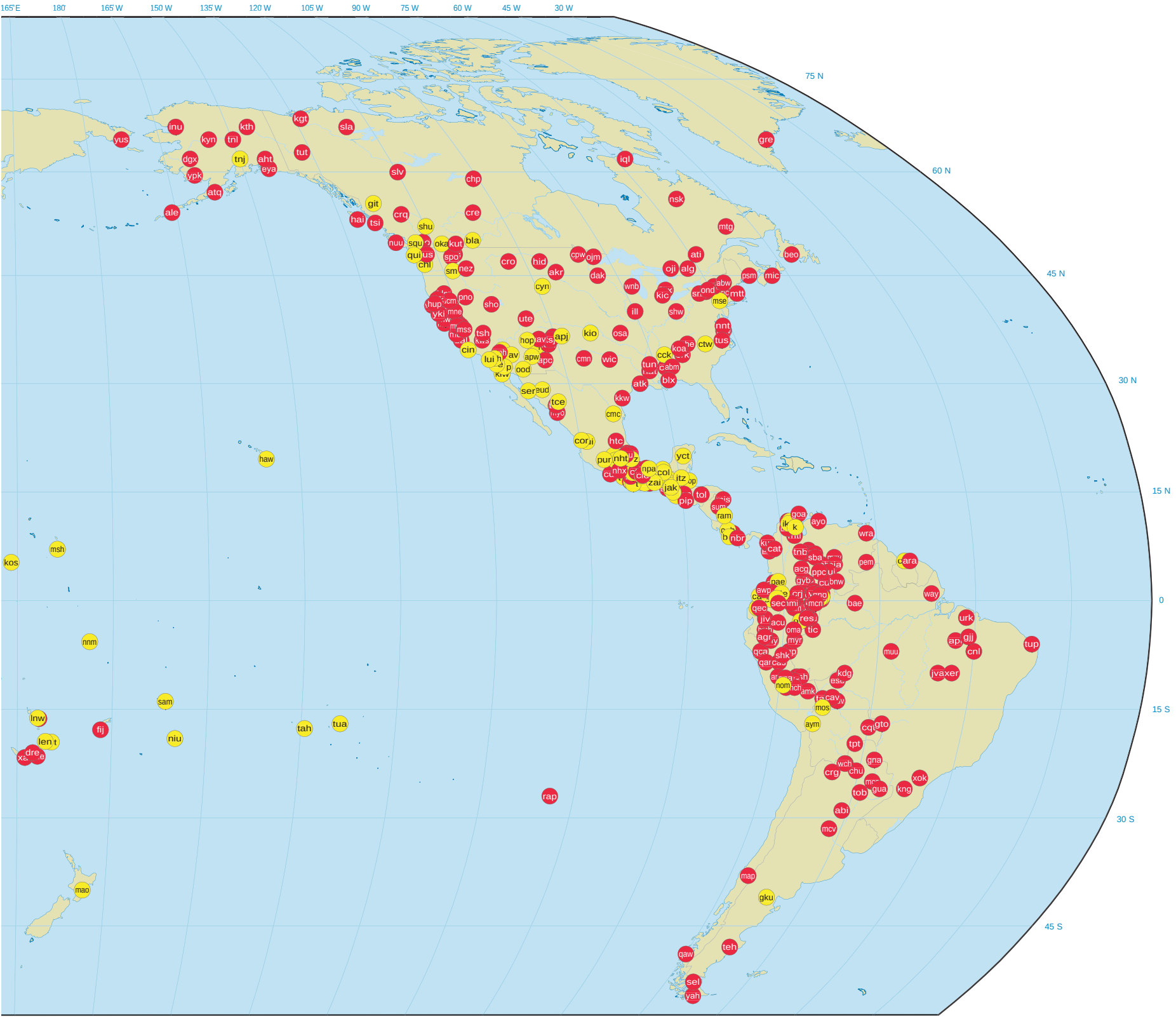




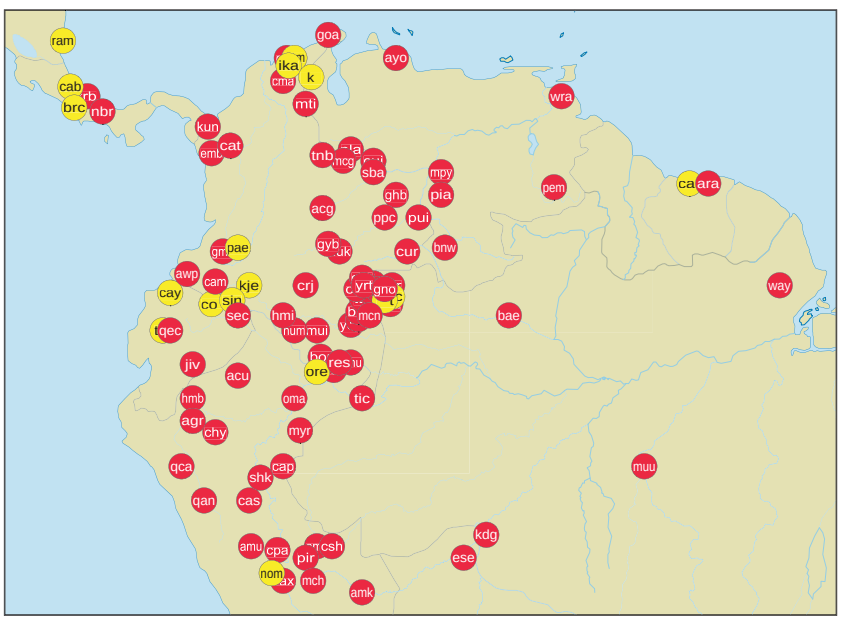
Vietnam and Philippines



western United States



Mesoamerica



northern South America

130 Finger and Hand

CECIL H. BROWN

1 Defining the values

This map shows the distribution of the two primary ways in which languages lexically treat the human finger and the hand of which it is a constituent. 'Finger' refers to any one of the hand's five appendages. ('Fingers' is a collection of two or more of the latter.) 'Hand' is defined as that part of the upper limb from the fingertips to the wrist. Two values are represented:

1.	Identity: a single word denotes both 'hand' and 'finger' and/or 'fingers'	72
2.	Differentiation: one word denotes 'hand' and another, different word denotes 'finger' (or, very rarely, 'fingers')	521
	total	593

English is an example of a type two language, with *finger* and *hand*. Another example is West-Central Oromo (Cushitic; Ethiopia and Kenya), with *quba* 'finger' and *harka* 'hand'.

In San Andrés Tzotzil (Mayan; Chiapas, Mexico), *c'obil* denotes 'hand' and an overtly marked construction based on the latter term, *bic'tal c'obil*, refers to 'finger'. Tzotzil is considered a differentiating language since the base word, *c'obil*, and the overtly marked construction, *bic'tal c'obil* (literally, 'little hand'), are different terms (although obviously nomenclaturally related). Similar examples of type two languages are Yapese (Austronesian; Micronesia), with *paaq* 'hand' and *bugul ii paaq* 'finger' (literally, 'tip of hand'), and Choctaw (Muskogean; Mississippi), with *ibbak* 'hand' and *ibbak ushi* 'finger' (literally, 'son of hand') (see Brown and Witkowski 1981 for further discussion). Designation of 'finger' through use of an overt marking construction based on a word for 'hand' is common among the differentiating languages sampled for the map. (In no language among those sampled is there an overtly marked construction for 'hand' based on a word for 'finger'.)

Examples of type one (identity) come from Warlpiri (Pama-Nyungan; Northern Territory, Australia), Cahuilla (Uto-Aztecan; California), and Kxoe (Khoisan; Namibia), respectively showing the terms *rdaka*, *-ma-l*, and *cèú*, all of which denote both 'hand' and 'finger'. Paakantyi (Pama-Nyungan; New South Wales, Australia), Mesa Grande Diegueño (Yuman; California), and Pacoh (Mon-Khmer; Vietnam and Laos) provide other examples of identity, with the respective terms *mara*, *esally*, and *ati*, all of which designate both 'hand' and 'fingers' (as a collection). Yugumbir Bandjalang (Pama-Nyungan; New South Wales, Australia) shows three different terms, each of which in addition to 'hand' is used to refer to the finger as a singular object or to the fingers as a collection.

A few languages included on the map show identity but nonetheless have a term for 'hand' that does not also designate 'finger'. For example, Oneida (Iroquoian; New York State) uses *osnúhsa* for both 'finger' and 'hand', and has an alternate term, *óhtsya*, for 'hand' (which also denotes palm, but not 'finger'). The convention followed here is that if a language shows 'finger/hand' polysemy, identity is judged present, even if there are terms in its lexicon denoting 'hand' that are not referentially extended to 'finger' as well.

2 Geographical distribution

The language sample of the map indicates that differentiating languages are about seven times more common than languages with identity. Differentiating languages occur in abundance all over the inhabited parts of the world. Identity languages occur relatively frequently in two major areas of the world: Australia and North America (excluding central and southern Mexico). Identity languages occur next most frequently in South America, but they are not especially prevalent there. Outside these three areas, such languages occur only sporadically and at very low frequencies (in Africa, mainland South-East Asia, New Guinea, and Polynesian islands of the Pacific).

3 Cultural distribution

Seventy-two languages of the sample of 593 show identity. Of these, forty-six (64 per cent) are spoken by traditional hunter-gatherers, eighteen (25 per cent) are spoken by groups having a mixed economy of cultivation and foraging, and eight (11 per cent) are spoken by full-fledged agrarians. Close to 90 per cent of type one languages, then, are spoken by peoples who traditionally have not embraced or have only partially embraced an agricultural way of life. The languages associated with each of these three cultural categories are as follows:

Hunter-gatherers: Alyawarra, Arrernte (Mparntwe), Atakapa, Atikamekw, Bandjalang, Bandjalang (Waalubal), Bandjalang (Yugumbir), Beothuk, Bunuba, Cahuilla, Comecrudo, Diegueño (Mesa Grande), Djabugay, Djingili, Hupa, Iñupiaq, Kalispel, Karankawa, Karok, Kaurna, Kxoe, Maranungku, Miwok (Northern Sierra), Miwok (Plains), Murrinh-Patha, Ndjébbana, Nez Perce, Ngadjumaja, Ngiyambaa, Nunggubuyu, Ojibwe (Minnesota), Paakantyi, Pintupi, Pitjantjatjara, Quileute, Seri, Shasta, Tasmanian, Thompson, Wambaya, Warlpiri, Wintu, Xokleng, Yana, Yir Yiront, Yukulta

Farmer-foragers: Amahuaca, Apache (Western), Carijona, Chayahuita, Cocopa, Eudeve, Javae, Jivaro, Mohawk, Yukpa, Munduruku, Natchez, Oneida, Onondaga, Seneca, Tunica, Urubú-Kaapor, Winnebago

Full-fledged farmers: Hawaiian, Keresan (Santa Ana), Navajo, Ngizim, Pacoh, Tahitian, Tuamotuan, Yagaria.

The geographical distribution of these three cultural categories is shown on Map 130A.

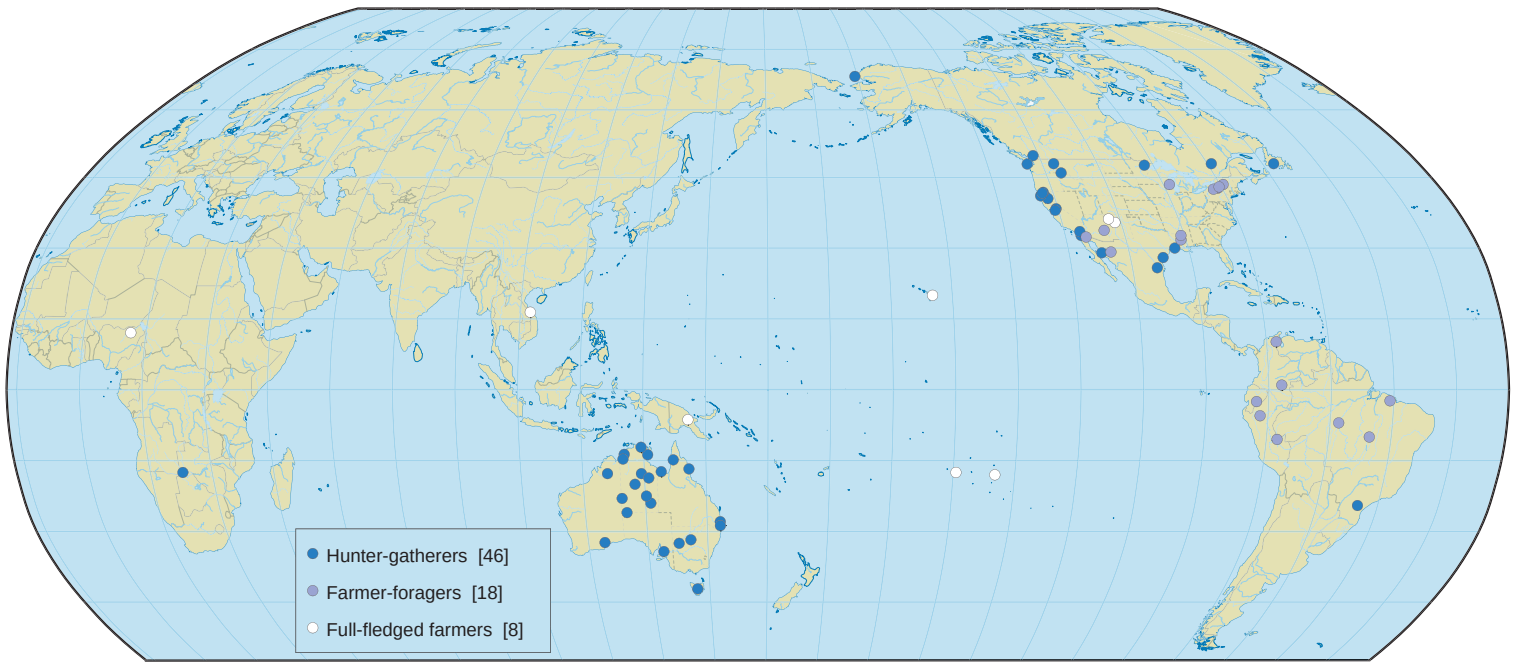
1.	Hunter-gatherers	46
2.	Farmer-foragers	18
3.	Full-fledged farmers	8
	total	72

Values of Map 130A Cultural Categories of Languages with Identity of 'Finger' and 'Hand'

The major concentration of languages showing identity in Australia and North America (excluding southern and central Mexico) reflects the fact that most of the hunter-gatherer languages sampled for the map are spoken in these two areas.

These data indicate that the way in which 'finger' and 'hand' are lexically distinguished is robustly associated with type of livelihood. Those languages spoken by groups who are not fully agrarian tend substantially more strongly to show identity than those of full agriculturalists. Conversely, languages of full-fledged farmers tend more strongly to show differentiation than those of hunter-gatherers and part-time farmers.

There are, of course, languages showing identity that nonetheless are spoken by full-fledged farmers (see above). In at least two such instances, evidence suggests that speakers of the language have relatively recently made the transition from hunting-gathering or from partial farming to full agriculture. Speakers of Navajo (Athapaskan; Arizona) are intrusive into the United States South-West from



Map 130A Cultural Categories of Languages with Identity of 'Finger' and Hand'

sub-Arctic North America, probably beginning about seven hundred years ago or so. In contemporary times, their Athapaskan-language-speaking relatives in the sub-Arctic region are traditional hunter-gatherers, such as the Navajo themselves almost certainly were when they began their southerly migration. Russell G. Schuh (p. c.) reports that while Ngizim (Chadic; Nigeria) is spoken by full-time farmers, hunting at one time was very important. In 1969, when he was doing linguistic fieldwork among the group, there was no game left other than small animals such as hares and grasscutters, but at the time they nonetheless talked considerably about hunting. In addition, the mainstay songs of the most famous Ngizim musicians are about hunters.

4 Explanatory framework

Farmers tend to lexically distinguish finger from hand more often than hunter-gatherers. This indicates that agriculturalists typically have more reasons for referring precisely to the finger than do foragers and, consequently, that the finger typically is more distinct for the former than for the latter. What, then, renders the finger more distinct among farmers compared to hunter-gatherers? Of several possible answers considered, the following is the one that, provisionally at least, seems most compelling.

Hunter-gatherer groups may typically differ from agrarians in the extent to which they make use of finger adornment. The major, globally spread, form of finger adornment is the use of rings. Plausibly, the manufacture and use of rings would enhance the salience of the finger as a distinct hand part and would serve to augment the number of contexts in which the finger is specifically referred to, thus promoting a term for finger different from that for hand. Hunter-gatherers traditionally may tend not to make and use finger rings to the degree farmers do. If so, this could figure prominently in an explanation of why foragers tend not to lexically distinguish finger from hand while agrarians do.

To my knowledge, no one has published a comprehensive survey of the distribution of finger rings across the world's peoples and cultures. Consequently, we do not know through rigorous comparative investigation that foragers tend to show less interest in finger rings than farmers. However, since agriculturalists generally have material culture inventories that are significantly larger than those of hunter-gatherers, plausibly, rings should be more prevalent in their inventories than in those of foragers.

The only relatively thorough source treating anthropological and historical aspects of rings is Kunz (1917). While anecdotal in nature, the information included in this work seems suggestive of the general absence of finger rings among foraging peoples. For example, Kunz (1917: 31) notes that rings are not favoured by the Eskimos, who neither make nor wear them. Indeed, he reports that Admiral Peary "found it impossible to dispose of a lot of rings he had taken with him on one of his Arctic trips in the belief that they would be attractive to the Eskimos, and good objects of barter." Kunz (1917: 17–18) also refers to archaeological excavations in various parts of the Americas (many in areas traditionally inhabited by hunter-gatherers and farmer-foragers) that have yielded no examples of prehistoric finger rings. He writes, "The attainable evidence in regard to the wearing of rings by aborigines of North and South America is, in the main, negative" (1917: 18).

I have undertaken an unsystematic electronic search of the Human Relations Area File database for mention of rings (for the finger) in a large number of published ethnographies. While references to finger rings are found in many monographs describing agricultural peoples, they are not common in those describing hunter-gatherer groups. When such references are found for foragers, they almost always involve rings that have been acquired by groups (usually from Western sources) rather than those natively manufactured.

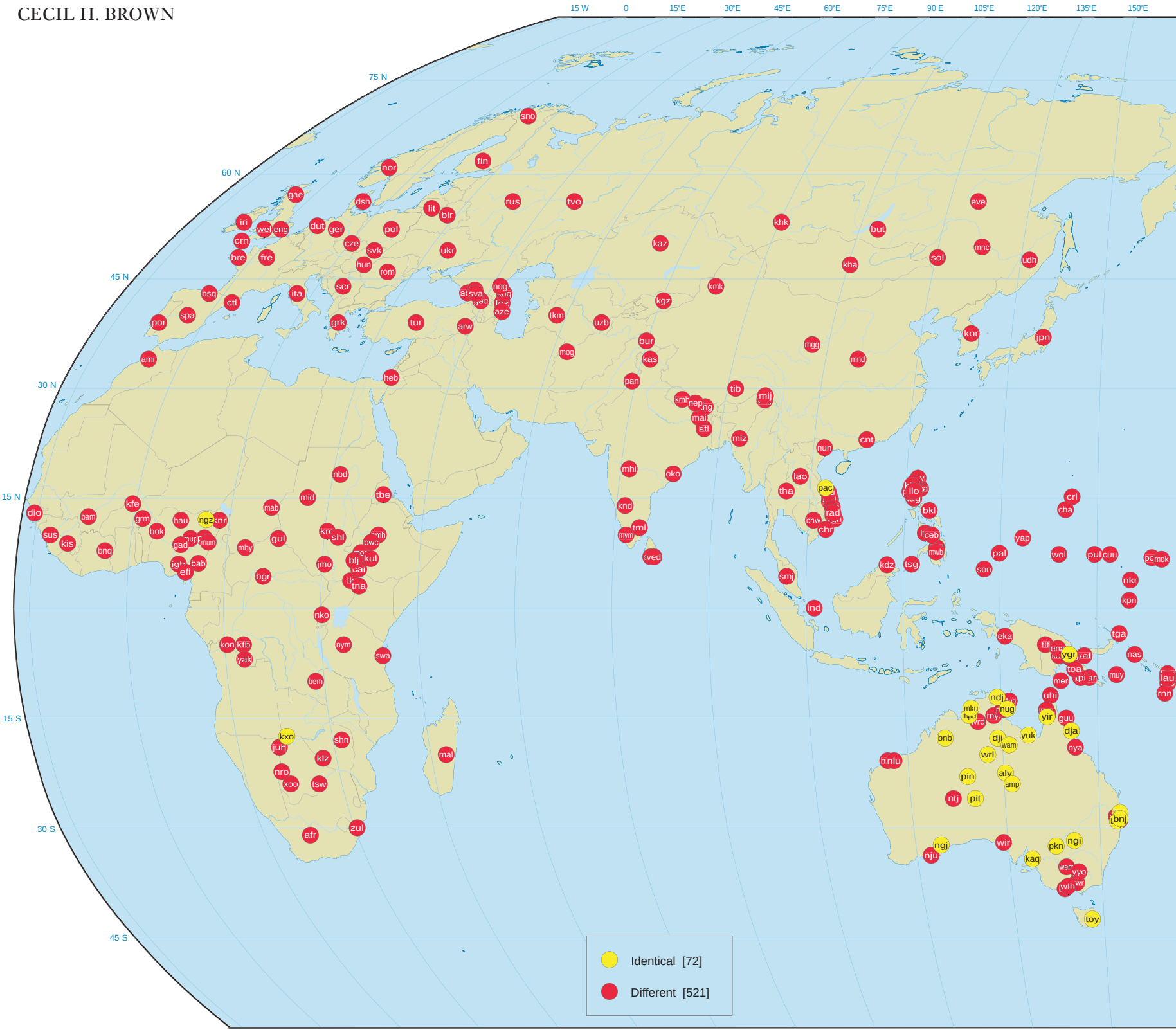
At present, then, the small amount of evidence assembled, mostly anecdotal in nature, supports the proposition that agricultural people typically make and use finger rings to a far greater extent than hunter-gatherers. This helps to explain why languages of farmers tend more strongly to lexically distinguish finger from hand than those of hunter-gatherers, which tend more strongly to use a single term to denote both 'finger' and 'hand'.

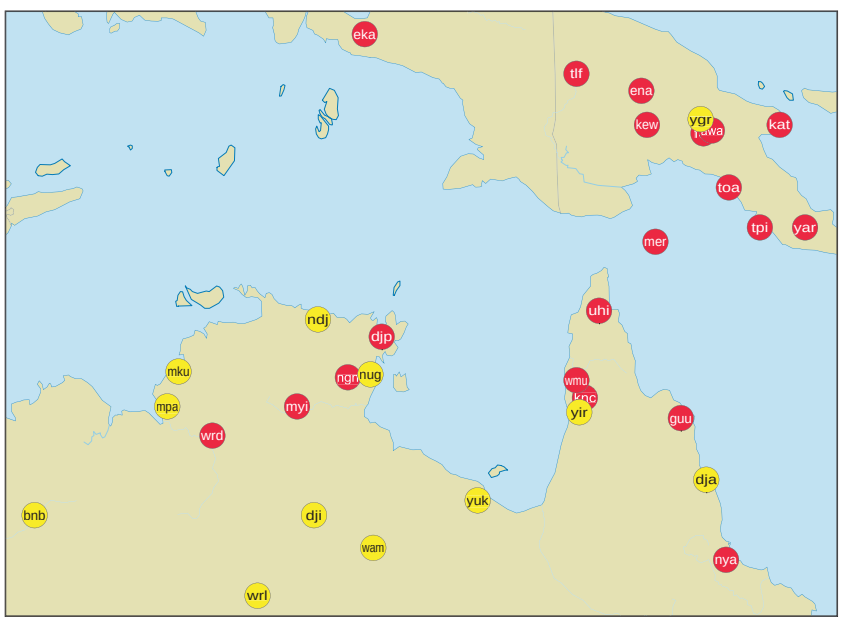
Acknowledgements

I am grateful to Giovanni Bennardo for suggesting the line of explanation in §4. Also, I would like to thank William Balée for supplying information on languages of South America and Russell G. Schuh for data on Ngizim speakers.

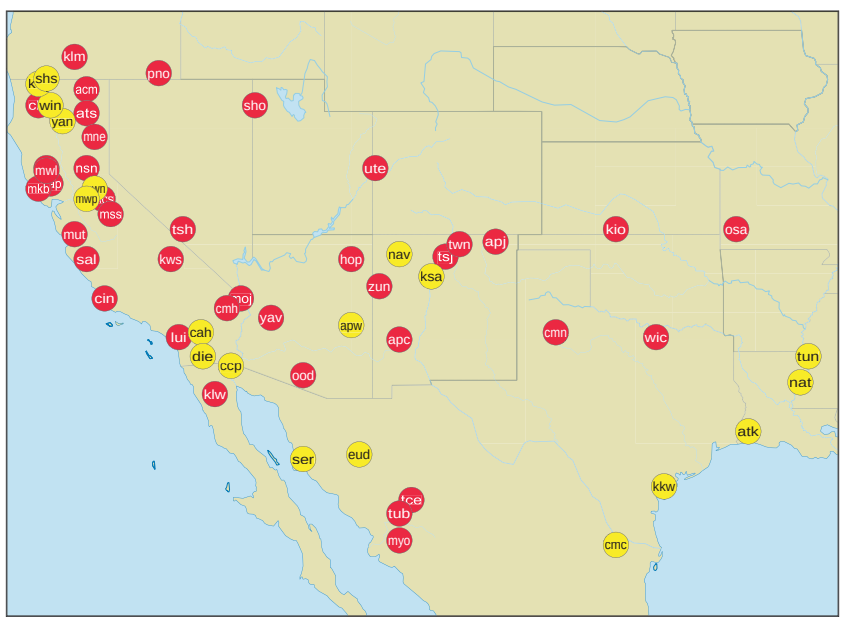
130 Finger and Hand

CECIL H. BROWN

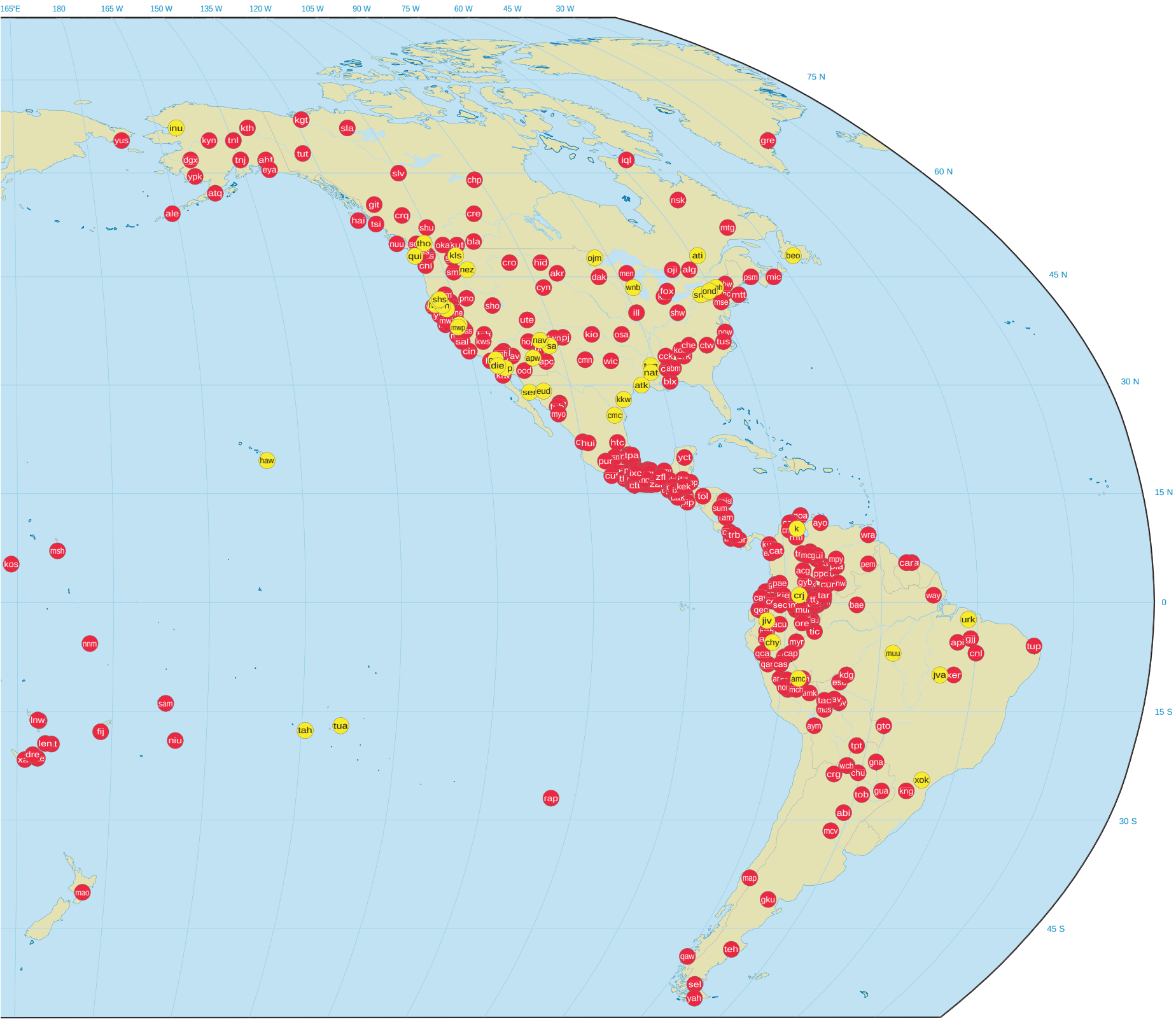


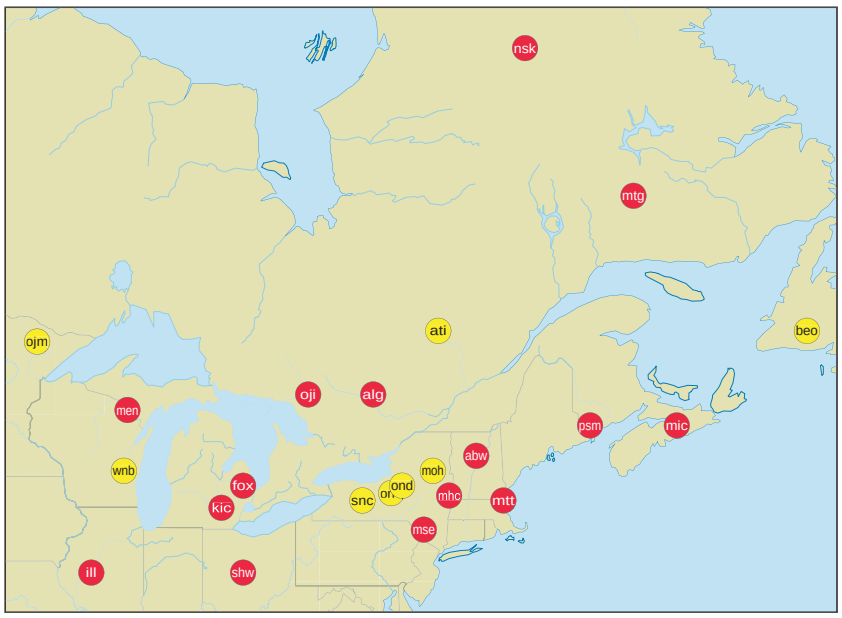


New Guinea and northern Australia

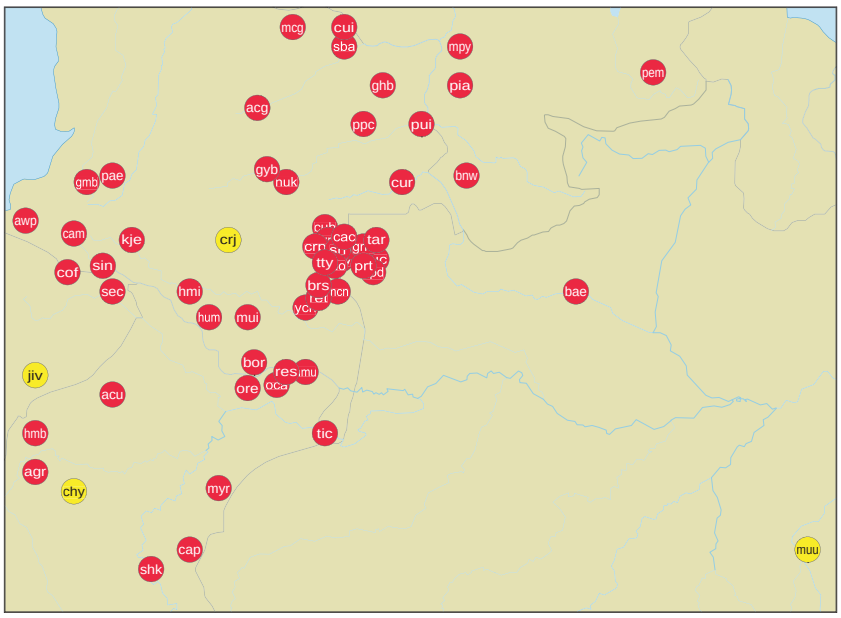


southwestern United States and northern Mexico





eastern North America



northwestern South America

131 Numeral Bases

BERNARD COMRIE

1 Introduction

This map is concerned with one aspect of the mathematical structure of linguistic expressions for numerals, namely the arithmetic base that is used in constructing numeral expressions. By the "**base**" of a numeral system we mean the value *n* such that numeral expressions are constructed according to the pattern . . . $xn + y$, i.e. some numeral *x* multiplied by the base plus some other numeral. (The order of elements is irrelevant, as are the particular conventions used in individual languages to indicate multiplication and addition.) A simple example is provided by Mandarin, with base ten, in which the numeral 26 is expressed as in (1).

(1) Mandarin
èr-shí-liù
two-ten-six

In Mandarin, the convention is that the numeral before the word for ten is to be multiplied by ten, while that after the word for ten is to be added to this product ($[2 \times 10] + 6$). Using this concept of base, plus some additional concepts to be introduced below, six main numeral systems can be identified, of which the second and third in the feature-value box can be viewed as subtypes of one superordinate type.

1. Decimal	125
2. Hybrid vigesimal-decimal	22
3. Pure vigesimal	20
4. Other base	5
5. Extended body-part system	4
6. Restricted	20
total	196

As the Mandarin example shows, the crucial concepts needed to demonstrate that a numeral system has a particular base are addition and multiplication. Beyond this, many numeral systems also make use of exponentiation of the base, i.e. expressions to denote the result of raising the base to various powers. Thus, English has a decimal system, and has a special term for 10^2, namely *hundred*, as well as one for 10^3, namely *thousand*. While the use of exponentiation often reinforces the identification of the base, it is not taken here as a defining feature, since some languages use addition and multiplication but without making use of exponentiation; an example is Chukchi (Chukotko-Kamchatkan; eastern Siberia), with a vigesimal (base twenty) system, but no special expression for 20^2, i.e. 400, which is simply expressed as 20×20. Moreover, the linguistic expression of exponentiation is often opaque—there is nothing in the form of the English words *hundred* and *thousand* to indicate that they are, respectively, the second and third powers of the base. Even the limited transparency provided by numerals like English *bi-llion*, *tri-llion*, with Latinate prefixes for two and three respectively, is only related by a quite complex formula to the corresponding power of the base 10: $10^{3(n+1)}$ in American usage or 10^{6n} in traditional British usage. For further general discussion, see Comrie (1997), Greenberg (1978), and Hurford (1975).

2 The six types

The **decimal** system has already been introduced by means of the Mandarin example (1); the general structure of numerals in a decimal system is $x10 + y$.

In a **pure vigesimal** system, the base is consistently twenty, i.e. the general formula for constructing numerals is $x20 + y$. An example is provided in (2) by Diola-Fogny (Atlantic, Niger-Congo; Senegal), in which the numeral 51 is expressed as 'two twenties and eleven'.

(2) Diola-Fogny (Sapir 1965: 84–5)

bukan	*ku-gaba*	*di*	*uɲɛn*	*di*	*b-ək͡pn*
twenty	CL6-two	and	ten	and	CL9-one

For practical reasons—in particular, the frequency of the type in the world's languages—it is useful to distinguish a hybrid vigesimal-decimal system in which the numbers up to 99 are expressed vigesimally, but the system then shifts to being decimal for the expression of the hundreds, so that one ends up with expressions of the type $x100 + y20 + z$; this is illustrated in (3) by the Basque expression for 256:

(3) Basque (Oroitz Jauregi, p.c.)

berr-eun	*eta*	*berr-ogei-ta-hama-sei*
two-hundred	and	two-twenty-and-ten-six

Bases other than 10 and 20 are also attested, albeit rarely, among the world's languages. Ekari (Trans-New Guinea; Papua, Indonesia) makes use of a base of sixty, as illustrated in the expression for 71 in (4); the base of sixty was also used in the ancient Near Eastern language Sumerian.

(4) Ekari (Drabbe 1952: 30)

èna	*ma*	*gàati*	*dàimita*	*mutò*
one	and	ten	and	sixty

Some languages of the world have numeral systems that do not make use of an arithmetic base. One such system is the **extended body-part system**, here illustrated by a discussion of Kobon (Madang, Trans-New Guinea), which is quite typical of a number of languages of Highland New Guinea. Languages like Kobon make use of further body parts to extend the system beyond the ten fingers. In Kobon specifically, the names of the following body parts (on the left-hand side of the body) are used in order to count from one to twelve: little finger, ring finger, middle finger, index finger, thumb, wrist, middle of forearm, inside of elbow, middle of upper arm, shoulder, collarbone, hole above breastbone. The count can then continue down the right-hand side of the body, from the collarbone to the (right) shoulder as 13 to the little finger as 23. It is then possible to reverse the count, starting from the little finger of the right hand as 24 back up to the hole above the breastbone as 35 and down again to the little finger of the left hand as 46. One effect of this is that the names of particular body parts when used as numerals are multiply ambiguous. For instance, *siduŋ* 'shoulder' can denote either 10 (on the left-hand side of the first pass across the body), or 14 (on the right-hand side of that pass), or 33 (on the right-hand side of the return pass across the body), or 37 (on the left-hand side of that pass), or 56 on the left-hand side of the next pass across the body, etc. There are usually means, optional or obligatory depending on the language, to distinguish the second side of the body used in a count from the first, as well as to indicate which pass across the body is being used, but there is no productive means to identify other than a small number of passes across the body. Extended body-part systems are thus typically rather limited in the range of numbers that they can express, but can be used productively at least into the scores.

Finally, some languages have **restricted** numeral systems, by which I mean more specifically a numeral system that does not effectively go above around twenty. The most restricted numeral system would of course be one lacking any numerals at all, and according to Dan Everett (p. c.) Pirahã (Mura; Brazil) is a language of just this type. A number of languages of the world have numeral systems that extend only as far as three (e.g. Mangarrayi (isolate; Northern Territory, Australia)), while others show slightly higher

but nonetheless heavily restricted upper limits, such as five (Yidiny (Pama-Nyungan; Queensland, Australia)).

3 Problem cases

In many cases, the assignment of a language to one or another of the types identified is straightforward, but nonetheless a number of problems can arise, and the following paragraphs will note some of these and, where relevant, indicate the solutions that have been adopted in the data analysis underlying this chapter.

First, it is essential to ascertain that the expressions in question are indeed numerals, since in many languages there are quantifying expressions, including some with quite specific denotations, other than numerals, such as *pair* in English (necessarily denoting a set of two). The general criterion to be used is that for an expression to qualify as a numeral, it must be the usual way of identifying that number of entities in the language in question within a noun phrase. In modern standard English, *seventy* (as in *seventy years*) is thus a numeral, whereas *three score and ten* is not, even if there may have been periods in the history of the English language, and may still be regional varieties, where the latter expression would qualify as a numeral. Note that it is probably not reasonable to require that numerals be used in counting—some cultures with low numeracy do not engage in counting, although they may nonetheless have a non-empty restricted numeral system.

Some languages have two (or more) numeral systems satisfying the criterion of the previous paragraph. Where both are of the same type, as with the indigenous and Sino-Korean systems in Korean—both decimal—then there is no problem in assigning the language unequivocally to one type. Where the systems are of different types, preference has been given to the most productive, e.g. the extended body-part system in Kobon, which exists alongside a restricted system.

Many languages combine different bases in the construction of their numeral system, and for the purposes of this chapter various decisions have been taken, some principled, some of a more practical nature, to limit the number of distinct types represented on the map. Only one mixed-base type has been given a separate representation, namely the type that is vigesimal in the range up to 99 and then decimal in the expression of the hundreds, because of its frequent occurrence among the languages of the world. In the case of other mixed systems, preference has been given to the base that is most productive in the construction of numerals in the range 20–400. In all numeral systems with a base of twenty or greater, and in several with a smaller base, numbers less than the base are constructed using smaller bases. For instance, in Igbo (Niger-Congo; Nigeria), with a base of twenty, the numerals 1–19 are constructed using ten as the base, as illustrated in (5) for 32:

(5) Igbo (Green and Igwe 1963: 37)

ohu	*nà*	*ìri*	*nà*	*àbụ̀ọ̀*
twenty	and	ten	and	two

In Supyire (Gur, Niger-Congo; Mali), with a base of eighty, the lower scores are expressed vigesimally, while numbers below twenty are expressed using a mixed quinary-decimal system (base five), as illustrated in (6), with 399 expressed as '$80 \times 4 + 20 \times 3 + 10 + 5 + 4$':

(6) Supyire (Carlson 1994: 169)

ỳ̀kwuu	*sicyɛɛré*	*'ná*	*béé-tàànrè*
eighty	four	and	twenty-three

ná	*ké*	*'ná*	*báárì-cyɛ̀ɛ̀rè*
and	ten	and	five-four

In some cases, an alternative base is used only in the construction of a small proportion of relevant numerals, and in such cases this alternative is simply disregarded in assigning the numeral system overall to a particular type. In French, for instance, the numerals in the range 80–99 have a vigesimal structure, as in the expression for 97 in (7), but since the system is otherwise entirely or almost entirely decimal, French has been assigned to the decimal type.

(7) French

quatre-vingt-dix-sept
four-twenty-ten-seven

While some languages, like Mandarin illustrated in (1), have completely or almost completely transparent structures of the type $xn + y$, other languages have various departures from the ideal formula in their numeral expressions through the appearance of morphophonological idiosyncrasies or portmanteau forms, and it is clearly necessary to be able to distinguish such exceptions from an instantiation of a different numeral system type. Relatively minor morphophonological idiosyncrasies can usually be identified without difficulty, as in relating English *fif-ty* to *five* and *ten*, with the identification of the first morpheme more transparent than that of the second. In some cases, completely different phonological forms may be used in certain combinations, as in the expression of the tens in Spanish, where the suffix *-enta*, as in *och-enta* 'eighty' (eight-ten), bears no formal resemblance to *diez* 'ten', but is nonetheless reasonably consistent in the expression of the tens. When only a handful of forms are portmanteau in an otherwise transparent system, they can be disregarded, as in the case of Russian monomorphemic *sorok* 'forty' in comparison with *pjat´-desjat* 'fifty' (five-ten). In some languages most or all of the products of the base are portmanteau forms, but we still identify them as such in that there is a separate such form for each product of the base and intermediate numbers are expressed by adding to that form; thus, in Turkish (Turkic; Turkey) each of the tens in the range 20–50 is monomorphemic (*yirmi* 'twenty', *otuz* 'thirty', *kırk* 'forty', *elli* 'fifty'; cf. *iki* 'two', *üç* 'three', *dört* 'four', *beş* 'five'), but numbers between the tens are expressed by addition of the remainder, as in (8), which expresses 21:

(8) Turkish (Kornfilt 1997: 428)

yirmi	*bir*
twenty	one

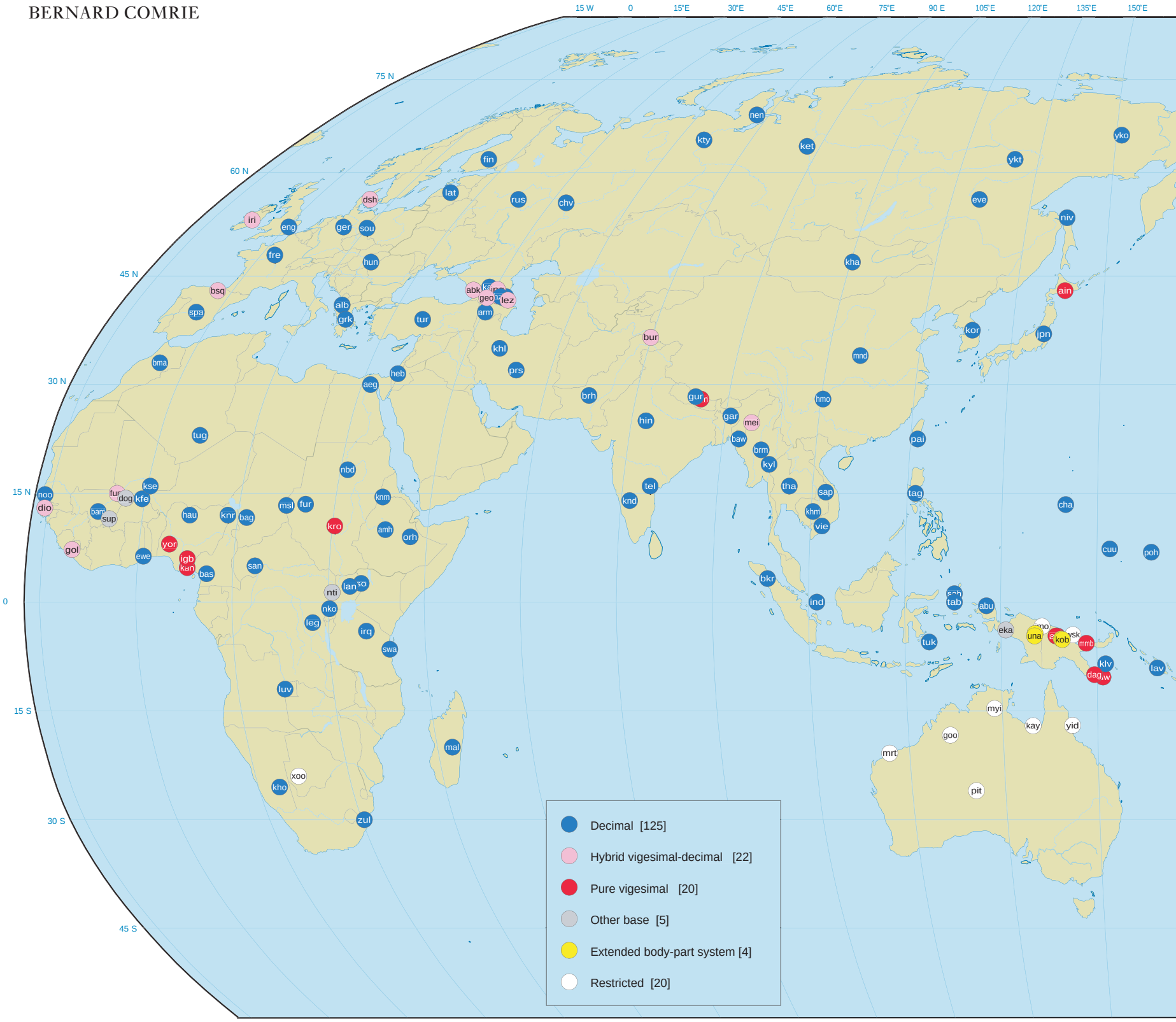
4 Geographical distribution

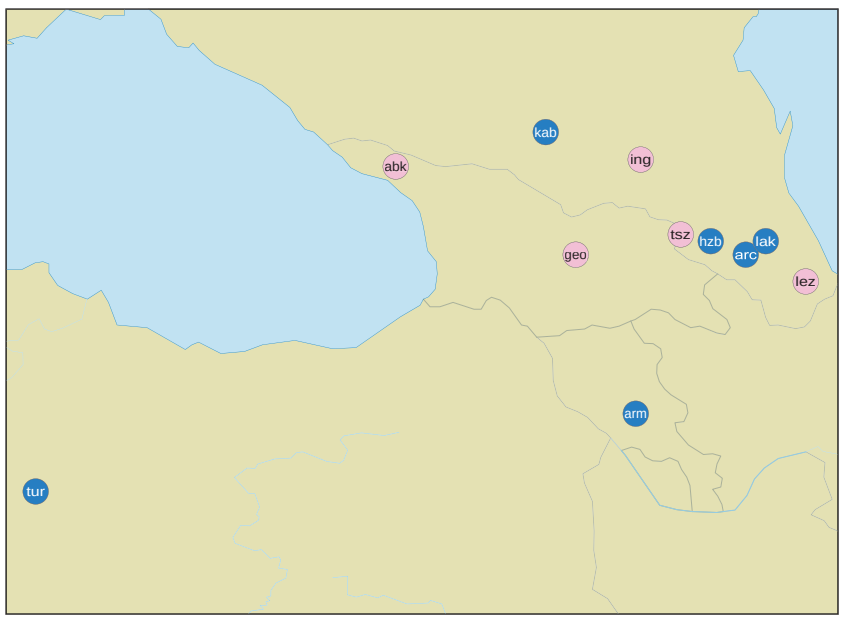
Even a cursory perusal of the accompanying map serves to show that, at least as far as numeral systems are concerned, we live in a decimal world, with the decimal type dominant in nearly every part of the world. Some of the other types are highly restricted geographically, in particular the extended body-part type, found as a basic numeral system in Highland New Guinea, and the restricted system, largely confined to Australia and Amazonia. Bases other than ten or twenty are extremely rare in the modern world, the examples in our sample being Supyire in West Africa and Ekari in Indonesian Papua.

However, the vigesimal system, whether pure or combined with the decimal system above 100, is still found in a number of different areas in the world, and is particularly frequent in some specific areas, such as Mesoamerica. Mesoamerica is in fact indicative of a worldwide historical trend for the dominant decimal system to encroach on and replace other systems. Pre-Conquest Mesoamerica was largely vigesimal, with the prototypical example being Classical Mayan. The influence of Spanish after the Conquest led to many indigenous languages adopting the Spanish numeral *ciento* 'one hundred' and with it the decimal system for the expression of the hundreds. In many languages, replacement by Spanish forms has percolated even further down the system, and it is not infrequent in contemporary accounts of Mesoamerican languages to read that in practice Spanish numeral expressions are used for all but the lowest numbers. Nondecimal numeral systems are even more endangered than the languages in which they occur.

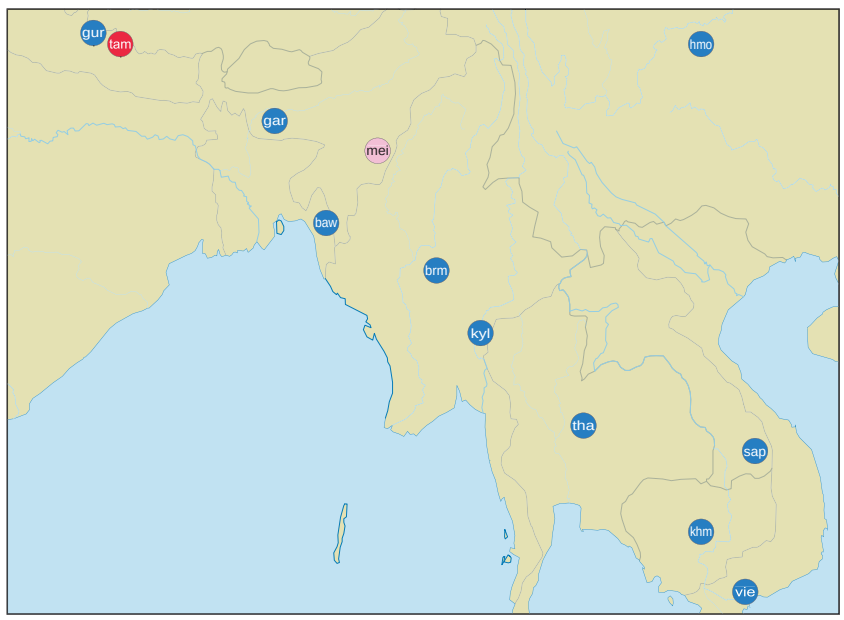
131 Numeral Bases

BERNARD COMRIE

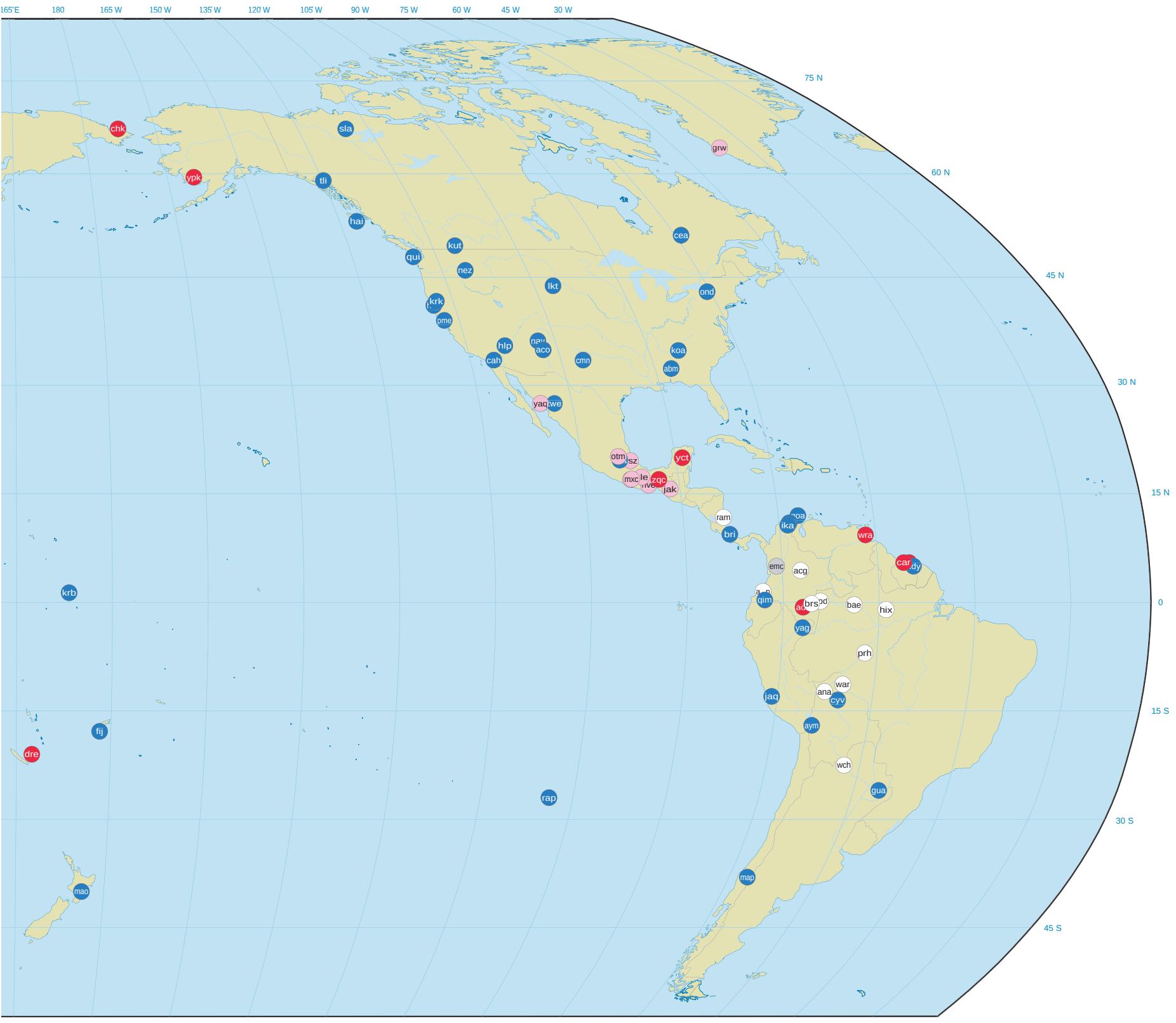


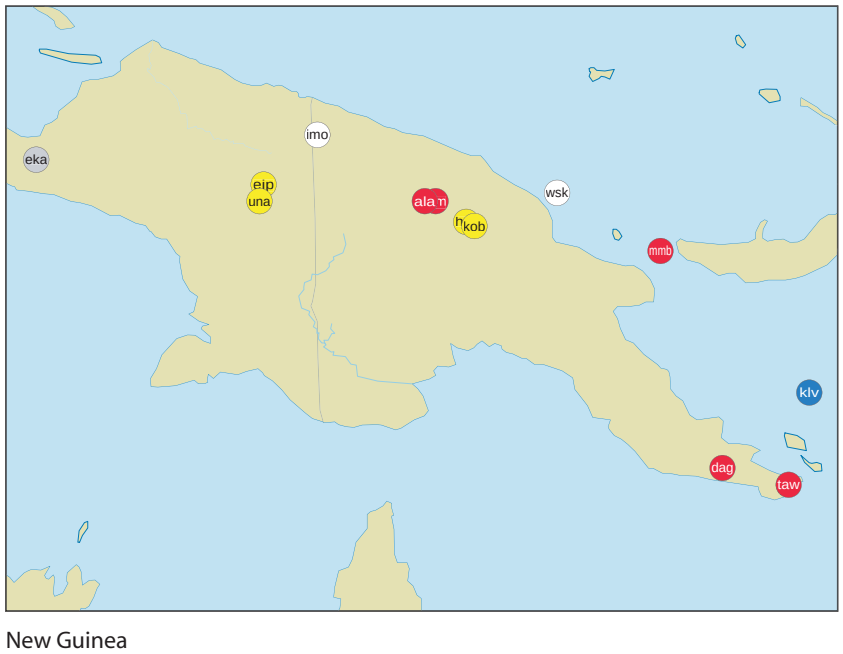


Caucasus and adjacent areas

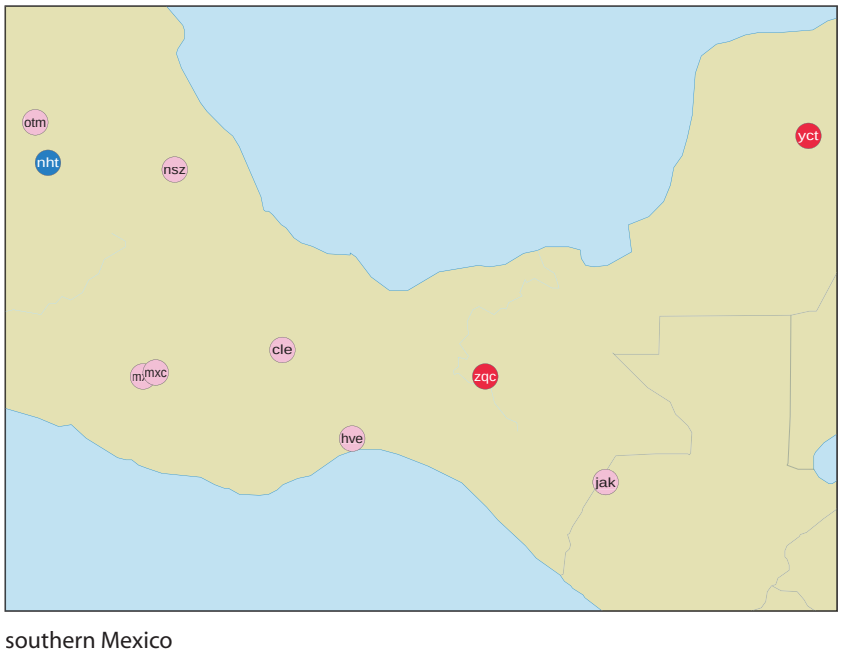


from Nepal to Vietnam





New Guinea



southern Mexico

132–135 Colour Terms

PAUL KAY AND LUISA MAFFI

1 The World Color Survey

The following four maps show the distribution of colour terms in some of the world's languages, based on the World Color Survey. This project, a collaboration of researchers at the University of California, Berkeley, and the Summer Institute of Linguistics (SIL), collected colour-naming data from 110 languages being studied by SIL linguist-translators in the late 1970s. For further details on methods and results of the World Color Survey see Kay et al. (1991), Kay and Berlin (1997), Kay et al. (1997), Hardin and Maffi (1997), Kay and Maffi (1999).

2 Basic colour term concepts

In recent models of the universals and evolution theory of basic colour terms, stemming originally from Berlin and Kay (1969), three kinds of **universal colour categories** are recognized (Kay and McDaniel 1978, Kay et al. 1991, Kay et al. 1997, Kay and Maffi 1999): primary colour categories, composite (disjunctive) colour categories, and derived (intersective) colour categories.

2.1 Primary colour categories. First, there are the six **primary colours** recognized in the widely accepted theory of **colour opponency** (Hering 1964 [1920], Jameson and Hurvich 1968; Hurvich 1982):

Black
White
Red
Yellow
Green
Blue

These are the two pairs of fundamental opponent chromatic colours <Red, Green> and <Yellow, Blue> and the single pair of fundamental achromatic colours <Black, White>. Black and white are not opposed in the same sense as the chromatic pairs. For example, we can see a combination of black and white (gray) but there is no sensation consisting of the combination of red and green or of blue and yellow. (That blue and yellow pigments can be mixed to produce a green pigment is not relevant; there are no bluish yellow or yellowish blue *sensations*.)

We will refer to the six colours black, white, red, yellow, green, and blue as **primaries**.

2.2 Composite colour categories. Secondly, there are **composite categories** consisting of fuzzy unions of the primaries. (Colour categories are modelled as fuzzy sets, that is, sets whose membership is a matter of more or less. For nontechnical discussion of fuzzy sets, fuzzy intersection, fuzzy union, and related concepts in relation to colour categories, see Kay and McDaniel (1978); for technical discussion of fuzzy sets and fuzzy logic, see Zadeh (1996).)

By far the most widespread composite in the ethnographic present is green-or-blue or "grue" (see Map 134). Over half the languages in the World Color Survey contain a grue term. Other major composite terms documented in the World Color Survey are, using "/" for fuzzy union, Black/Green/Blue, White/Red/Yellow, Black/Blue, Red/Yellow, Yellow/Green/Blue, and Yellow/Green. (There is one, somewhat unclear, instance of a probable White/Yellow term in Waorani, an isolate of Ecuador.)

There are also what we might call *minor composites*. For example, in systems that lack terms for purple and brown, these colours are often included in the black term. Similarly, in systems restricted to the primaries, nonprimary shades such as brown, purple, pink, orange, and gray are characteristically either split between two adjacent primaries or covered by a single adjacent primary.

2.3 Derived colour categories. The third type of widespread colour category are the derived (intersective) categories, which are experienced as mixtures of the primaries. These are based formally on fuzzy intersections of primaries. Gray is experienced as a mixture of Black and White, pink of White and Red, orange of Yellow and Red, purple of Blue and Red, and brown (arguably) of Yellow and Black.

3 An evolutionary sequence of basic colour systems

Figure 1 presents both a typology and an evolutionary sequence of basic colour systems covering 101 of the 110 World Color Survey languages. (The colour lexicons of the remaining nine languages are composed mostly of terms denoting categories appearing in Figure 1. Limits of space preclude further discussion here; see Kay and Maffi 1999.)

Basic colour lexicon types are shown in square brackets and evolutionary stages are represented by columns, labelled with Roman numerals. For example, stage I contains a single type, consisting of two terms, one a composite of White, Red and Yellow and the other

a composite of Black, Green, and Blue. (There are no stage I languages in the World Color Survey sample. Stage I is shown here for theoretical completeness.)

Arrows indicate observed evolutionary transitions between types. Arabic numerals within brackets indicate the number of World Color Survey languages conforming to the corresponding colour lexicon type. Arabic numerals placed between basic colour lexicon types indicate the number of World Color Survey languages transitional between the two types.

Derived categories are not considered in the typology presented in Figure 1. Usually, derived categories appear only after stage V, where each primary has received a separate term, but there are many exceptions. Creating a type for every resulting combination would lead to pointless complexity. Composite terms are characteristic of the earliest stages of development. Evolutionary development consists essentially in a succession of increasingly fine partitions of the perceptual colour space. The vast majority of World Color Survey languages, 83 per cent, belong to a type or transitional type along the single evolutionary trajectory I > II > $III_{Green/Blue}$ > $IV_{Green/Blue}$ > V.

4 The colour term maps

Four colour term maps are presented. They tabulate the data from the 110 World Color Survey languages and comparable data from nine additional languages (Berlin and Kay 1969; Xu 1994). The World Color Survey sample hardly overlaps with the standard sample of the *World Atlas of Language Structures* and care should be exercised in comparing the colour maps with other maps in this atlas.

The four colour term maps address four major types of questions a reader might have regarding the distribution of a basic colour term system in the world:

(i) What is the basic evolutionary stage? That is, how many nonderived (i.e. either composite or primary) terms does the system contain? This amounts to assigning each language to a column in Figure 1.

(ii) How many basic colour terms (including derived terms) does the system contain?

(iii) How are the "cool" primaries green and blue treated?

(iv) How are the "warm" primaries red and yellow treated?

4.1 Map 132 Number of Nonderived Basic Colour Categories. Because there is sometimes synonymy in basic colour terms, it is more revealing to tabulate the number of **categories** that are encoded by basic colour terms than to count the number of basic colour terms per se. The number of nonderived basic colour categories encoded in basic colour terms is a feature with self-explanatory integer values, with the exception that we define values two, four, and six to represent colour lexicons judged to be intermediate between two types which encode the adjacent whole numbers of nonderived basic colour categories.

1. 3 categories	10
2. Between 3 and 4 categories	3
3. 4 categories	9
4. Between 4 and 5 categories	1
5. 5 categories	56
6. Between 5 and 6 categories	11
7. 6 categories	29
total	119

The most frequent type in this sample of mainly unwritten contemporary languages is one with five nonderived basic colour categories encoded. Languages with fewer than five nonderived basic colour categories encoded as separate terms tend strongly to occur in the tropics.

4.2 Map 133 Number of Basic Colour Categories. This feature is similar to the preceding except that in this case (i) all basic colour categories, including derived categories, are counted; and (ii) *ranges* of numbers of categories form the values of the feature.

1. 3, between 3 and 4, or 4 categories	20
2. Between 4 and 5, 5, or between 5 and 6 categories	26
3. 6 or between 6 and 7 categories	34
4. 7 or between 7 and 8 categories	14
5. 8 or between 8 and 9 categories	6
6. 9, between 9 and 10, or 10 categories	8
7. More than 10 categories	11
total	119

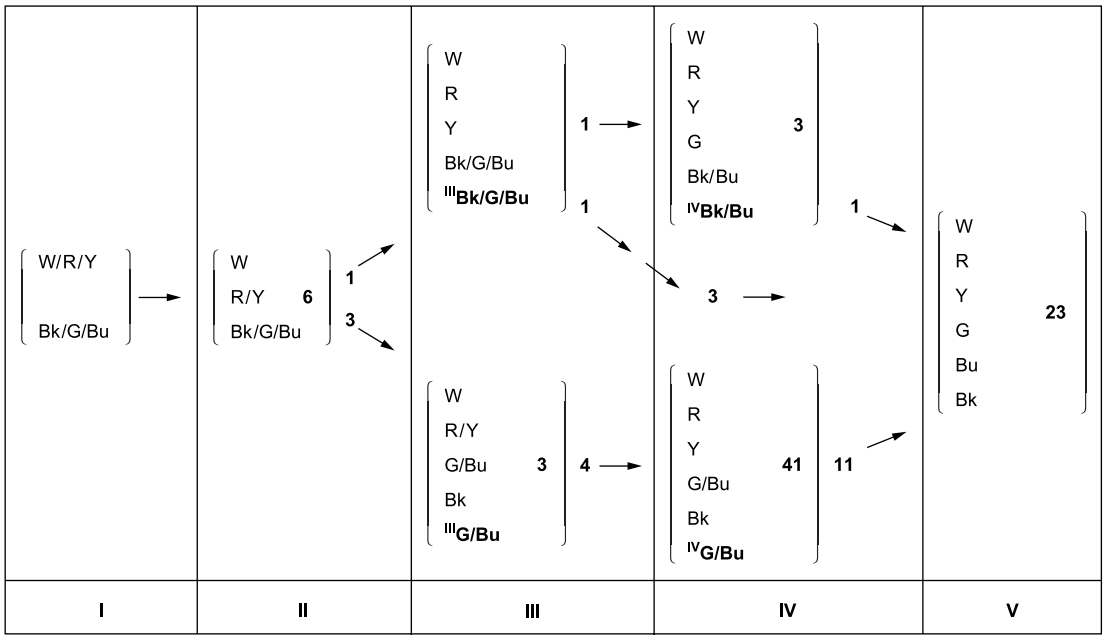


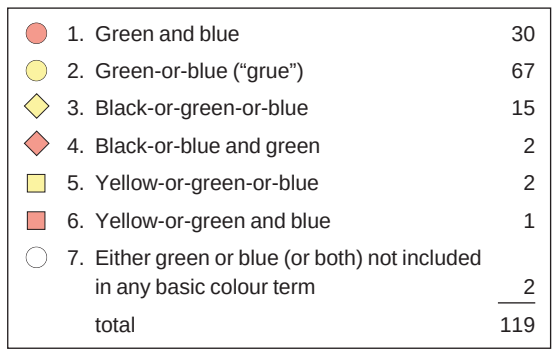
Figure 1 A Typological and Evolutionary Scheme Covering Most World Color Survey Languages
(Source: Kay and Maffi 1999, Figure 3.)

As on the previous map, a tendency may be observed for the number of basic colour categories encoded to vary inversely with distance from the equator.

4.3 Map 134 Green and Blue. Given the typology of Figure 1, and the fact (not noted there) that Green and Yellow can be associated, there are six ways green and blue can be encoded.

(i) Green and blue each get a basic colour term (as in English, for example).
(ii) A green-or-blue category gets a basic colour term ("grue").
(iii) A black-or-green-or-blue category gets a basic colour term.
(iv) A black-or-blue category gets a basic colour term and green gets a separate basic colour term.
(v) A yellow-or-green-or-blue category gets a basic colour term.
(vi) There is a term for yellow-or-green and a term for blue.

In addition, we need a seventh value, not applicable, for languages in which either the green sensation or the blue sensation (or both) fails to be included in any basic colour term.



The most frequent way of encoding green and blue in this sample of languages is in a grue term.

4.4 Map 135 Red and Yellow. Given the typology of Figure 1, and (as before) the fact that Green and Yellow can be associated, there are four ways red and yellow can be encoded—as well as a value "not applicable".

(i) Red and yellow each gets a basic colour term.

(ii) Red-or-yellow is a basic colour term.

(iii) Red gets a separate basic colour term and yellow-or-green-or-blue is encoded as a basic colour term.

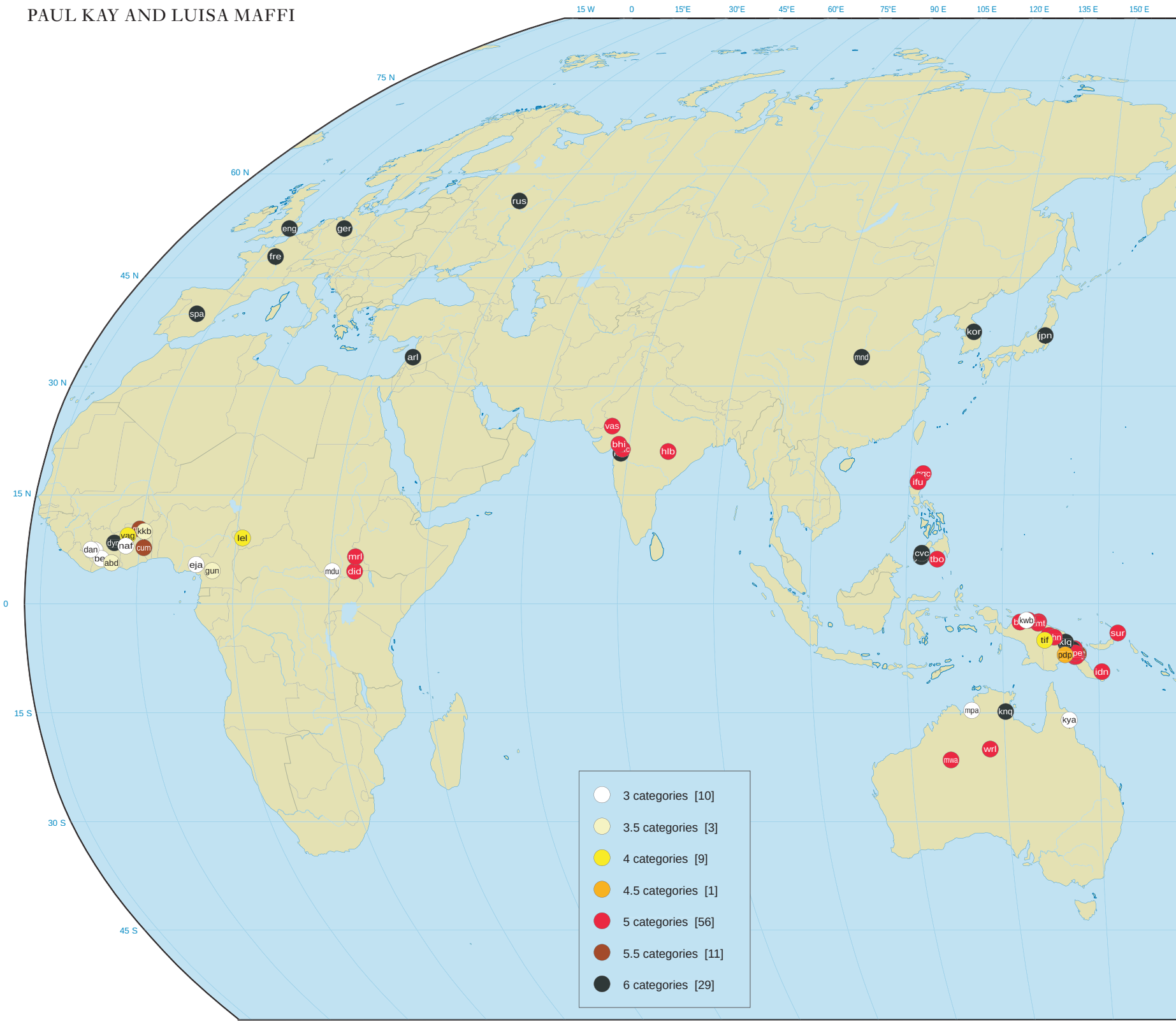
(iv) Red is a separate basic colour term and yellow-or-green is encoded as a basic colour term.

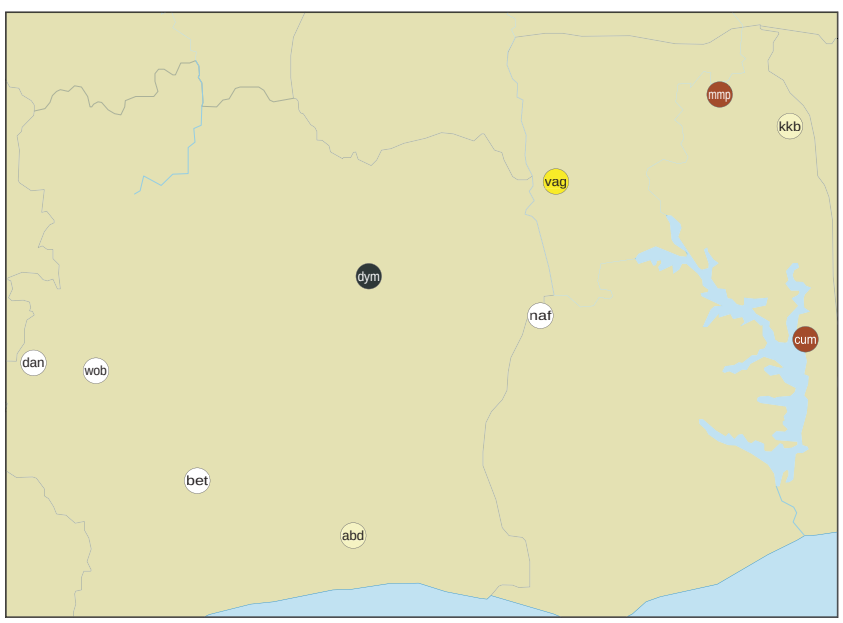
1.	Red and yellow	97
2.	Red-or-yellow	15
3.	Red and yellow-or-green-or-blue	3
4.	Red and yellow-or-green	1
5.	Either red or yellow (or both) not included in any basic colour term	3
	total	119

Red and yellow are much less frequently encoded in a single composite than green and blue. In fact, no language has to our knowledge been reported – in the World Color Survey or elsewhere—that gives separate basic colour terms to green and blue while retaining a red-or-yellow composite.

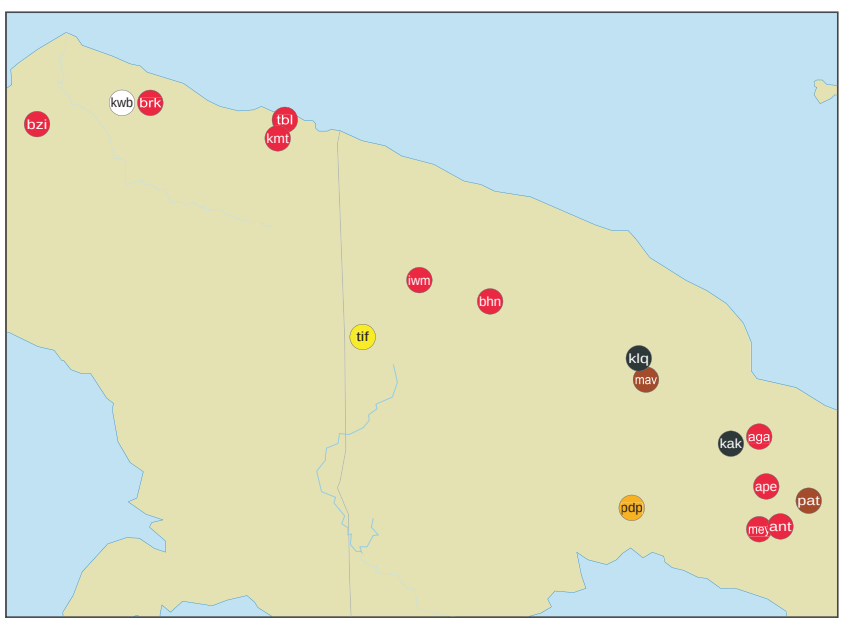
132 Number of Nonderived Basic Colour Categories

PAUL KAY AND LUISA MAFFI

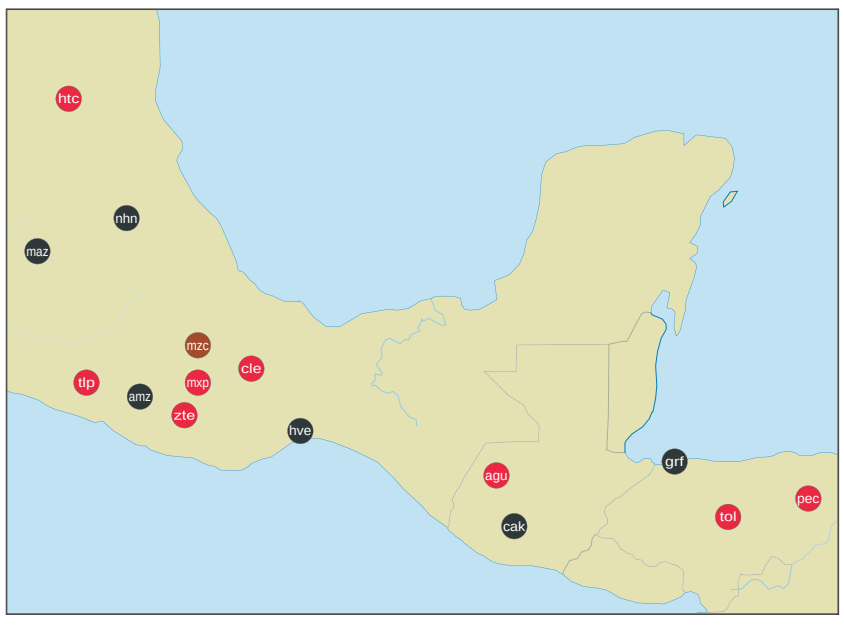




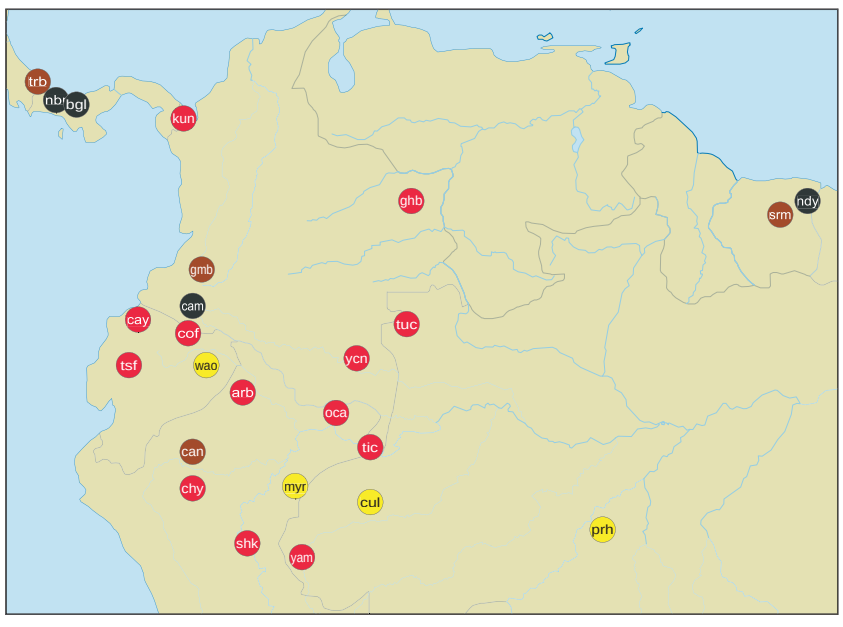
Côte d'Ivoire and Ghana



New Guinea



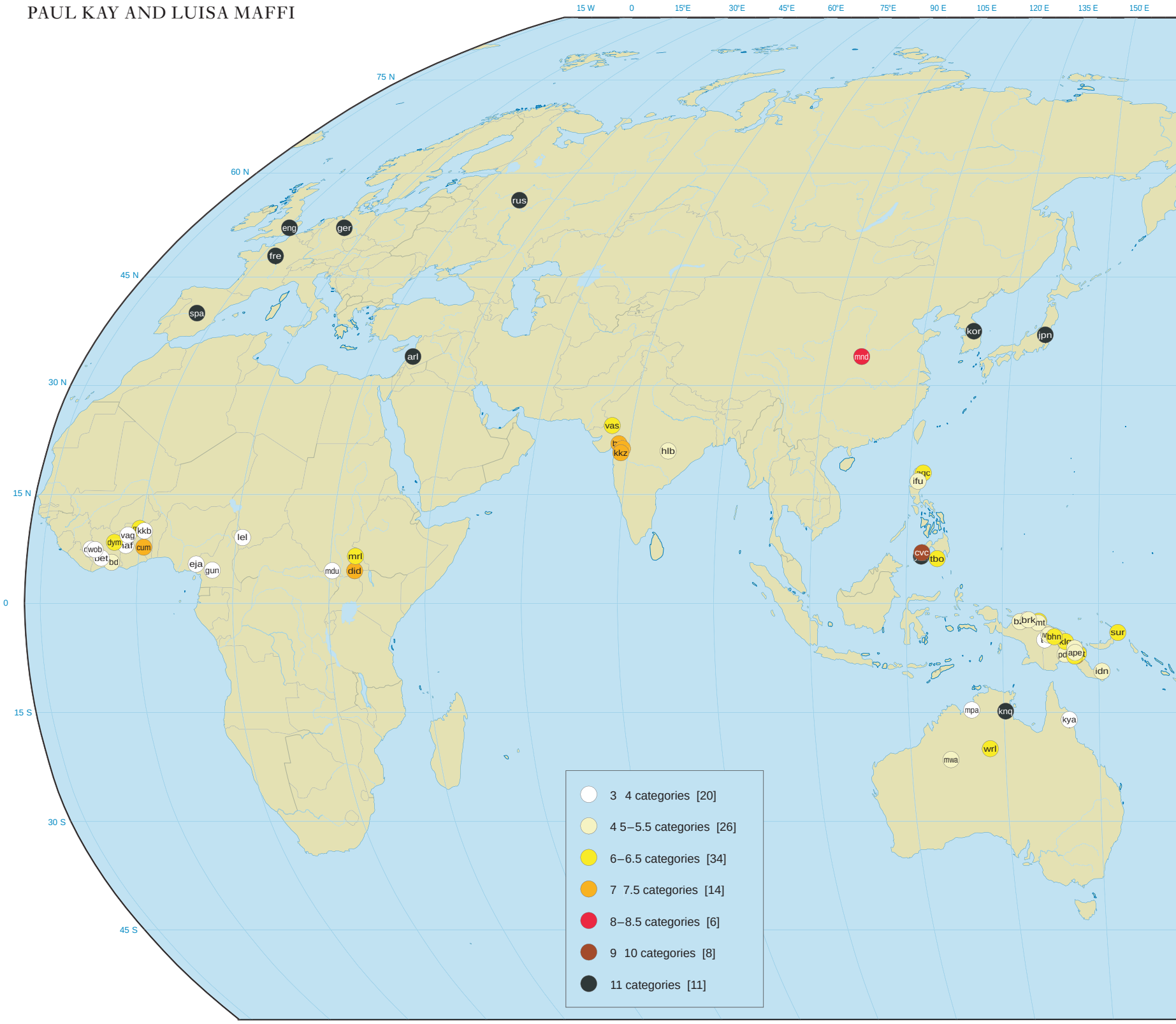
Mesoamerica

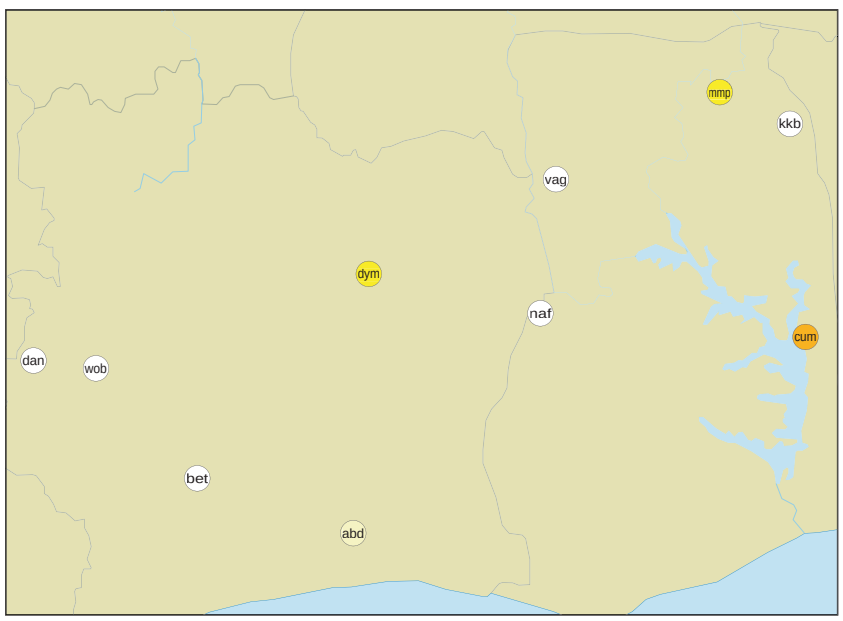


northwestern South America

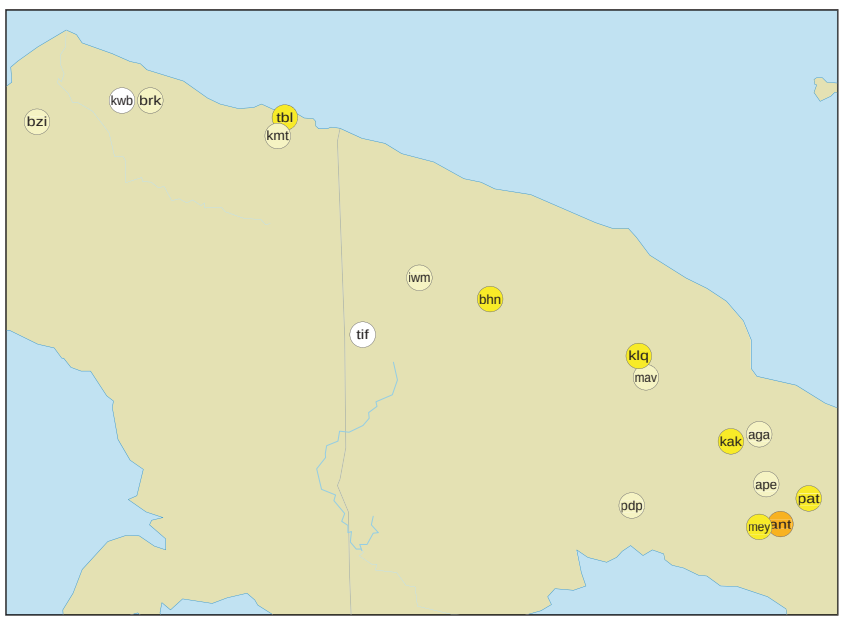
133 Number of Basic Colour Categories

PAUL KAY AND LUISA MAFFI

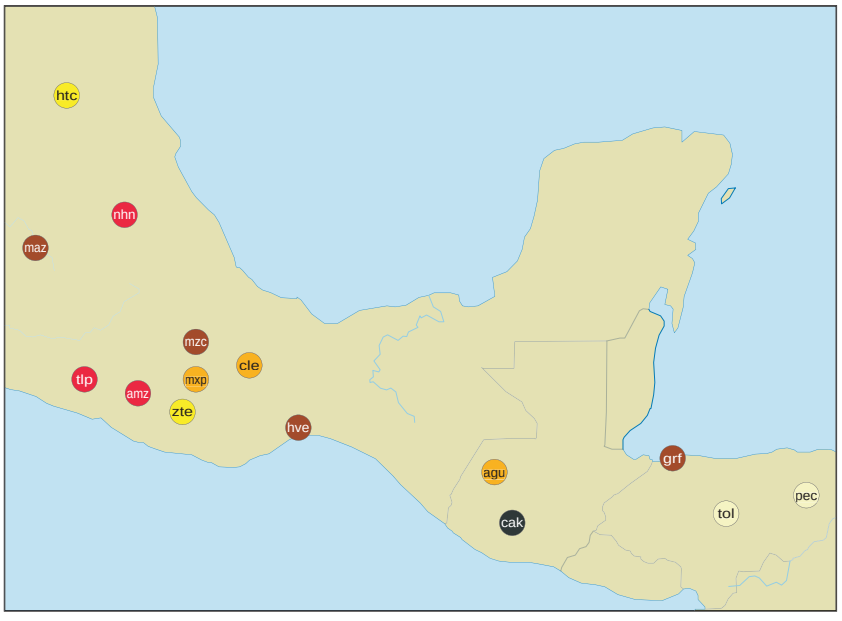




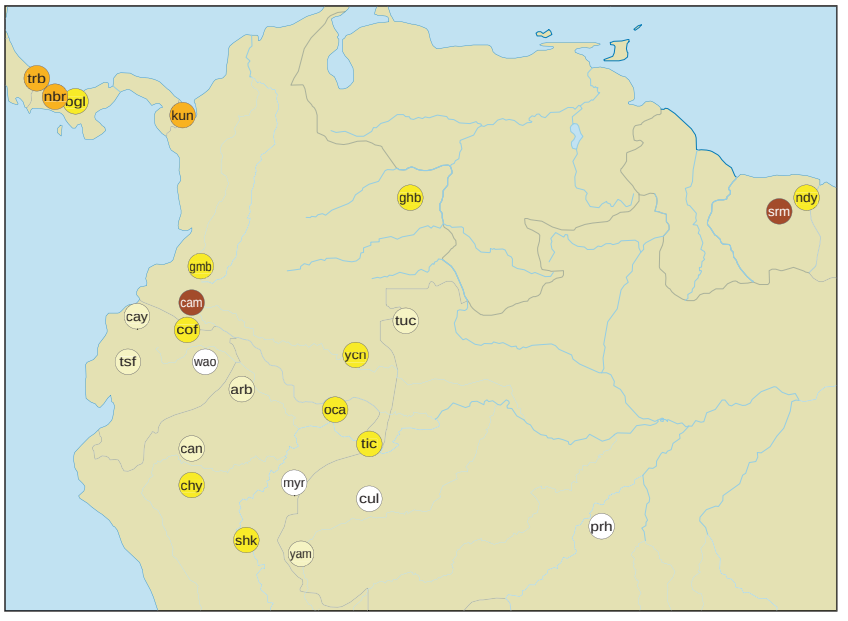
Côte d'Ivoire and Ghana



central New Guinea



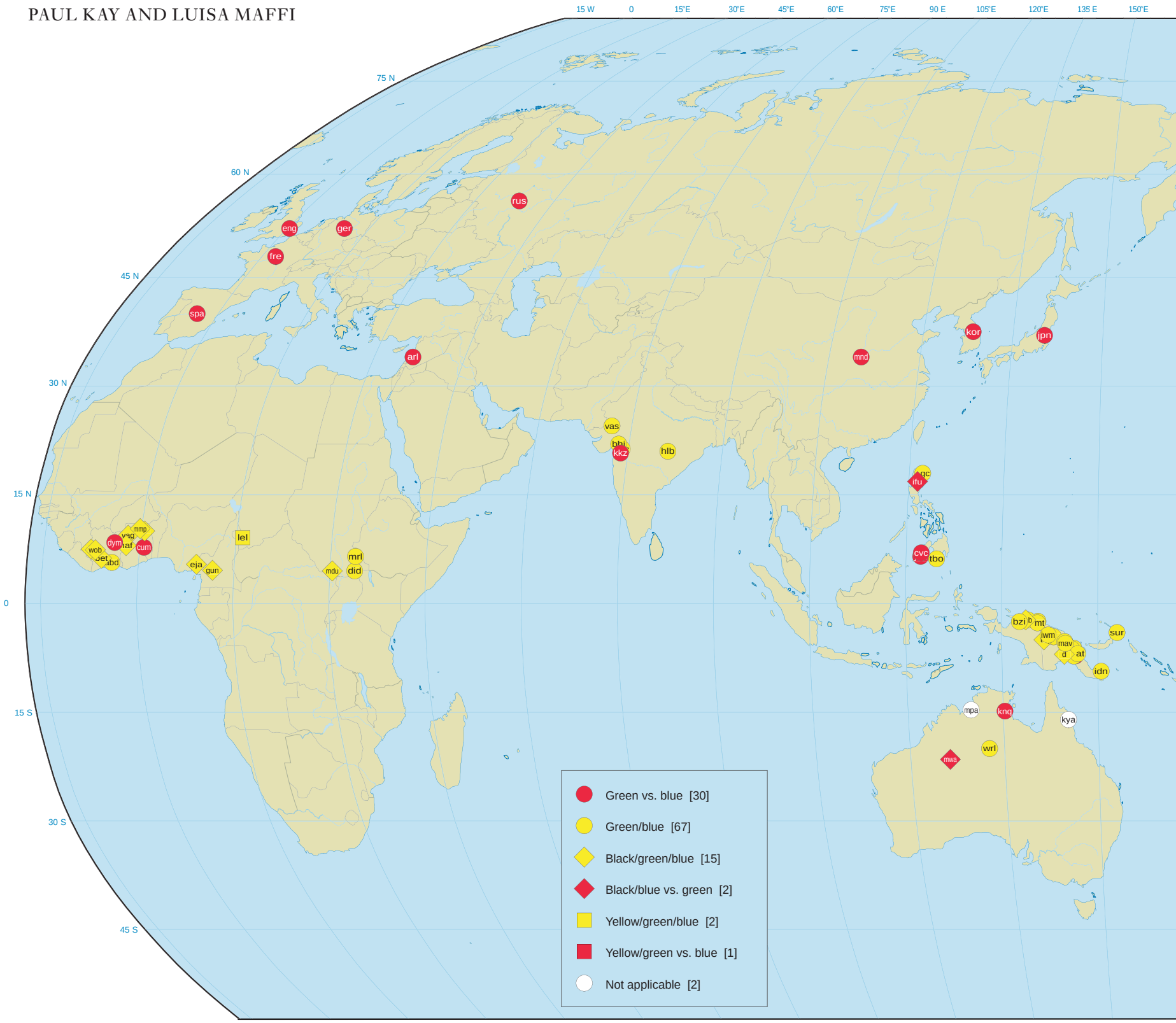
Mesoamerica

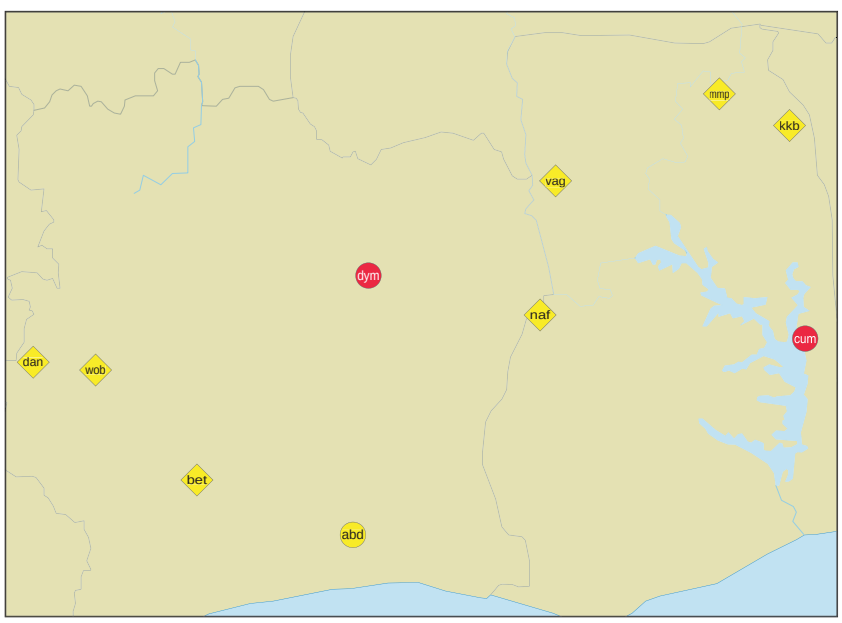


northwestern South America

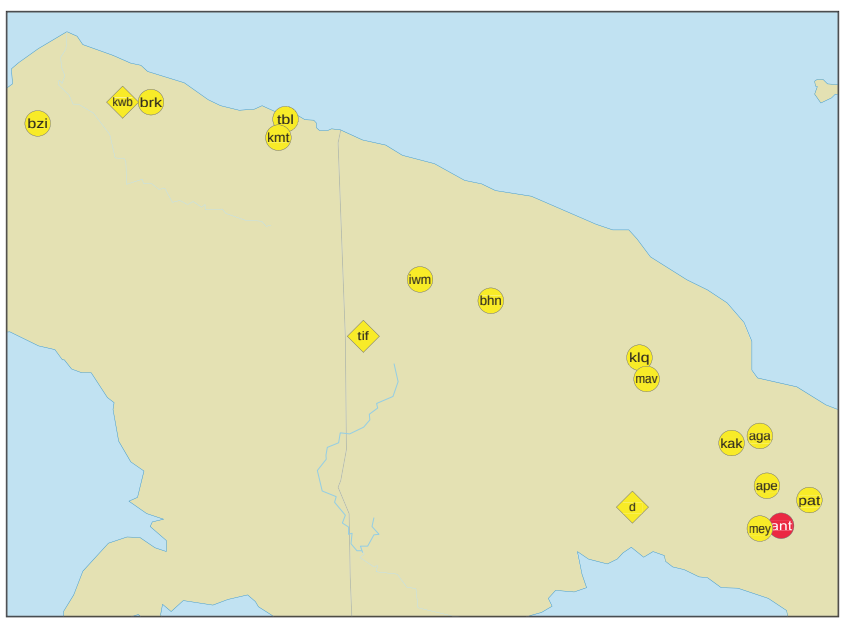
134 Green and Blue

PAUL KAY AND LUISA MAFFI

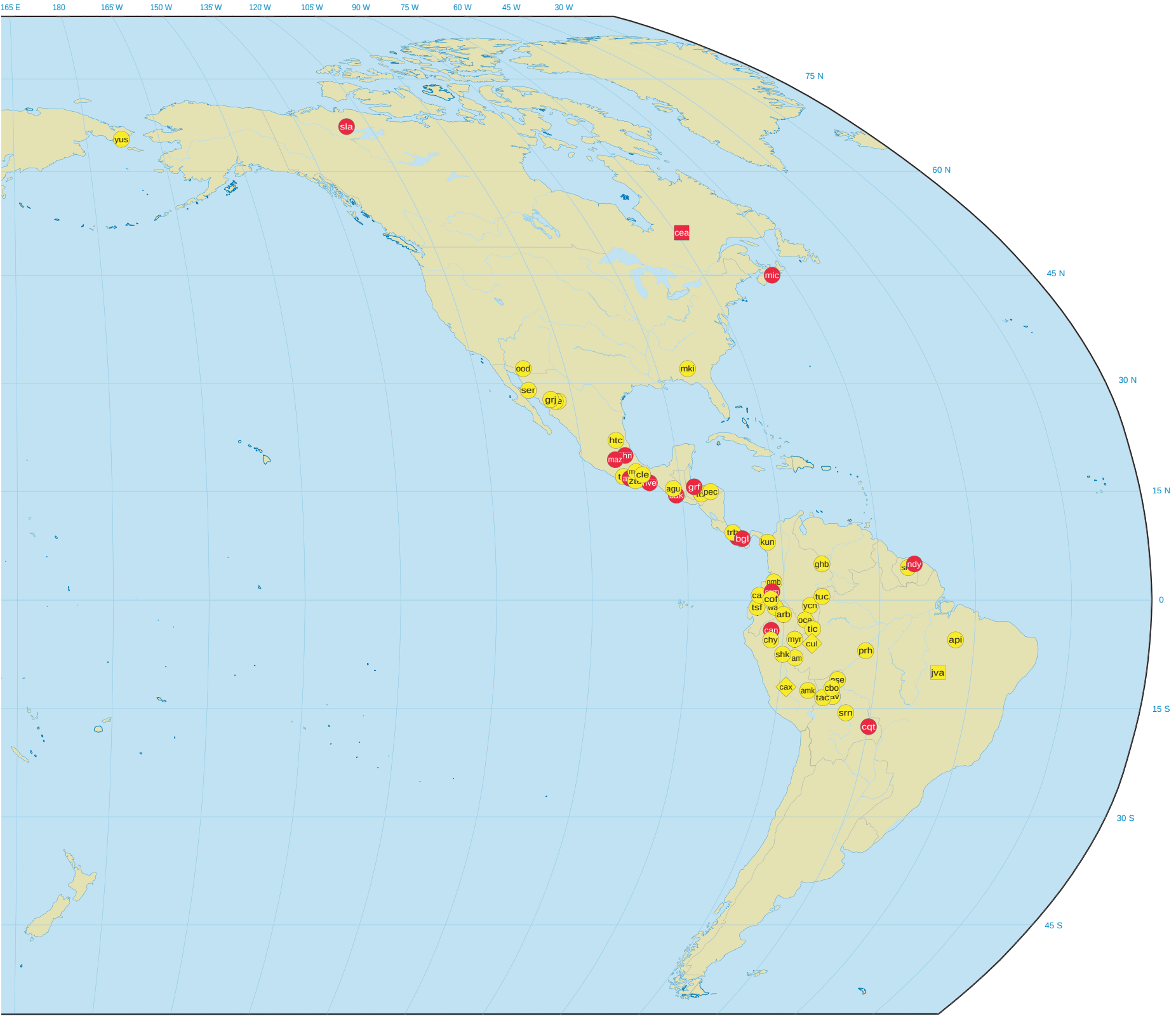


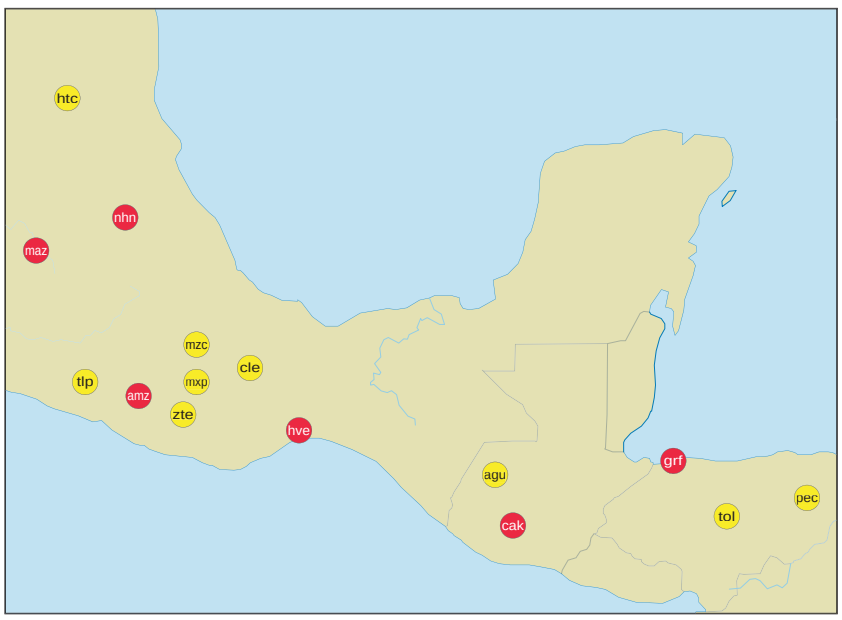


Côte d'Ivoire and Ghana

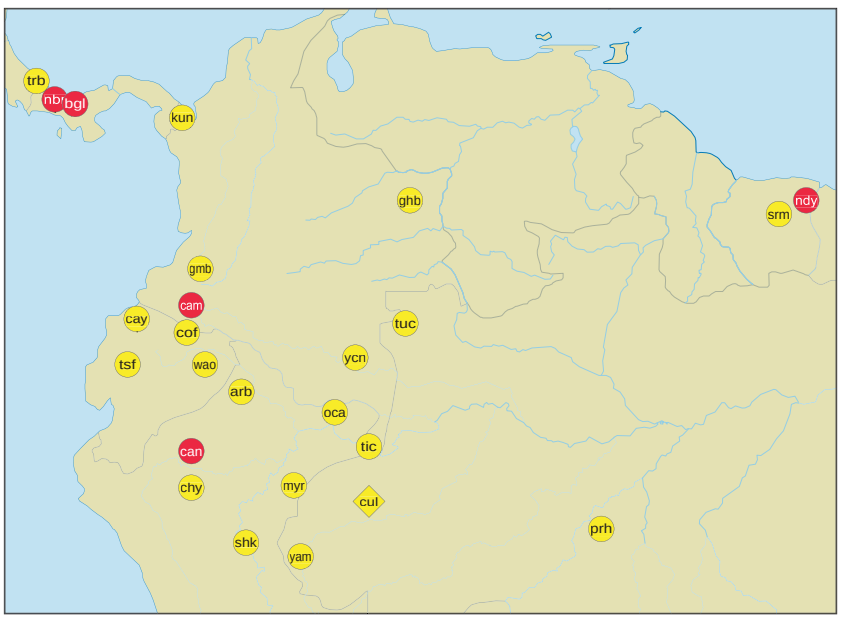


New Guinea





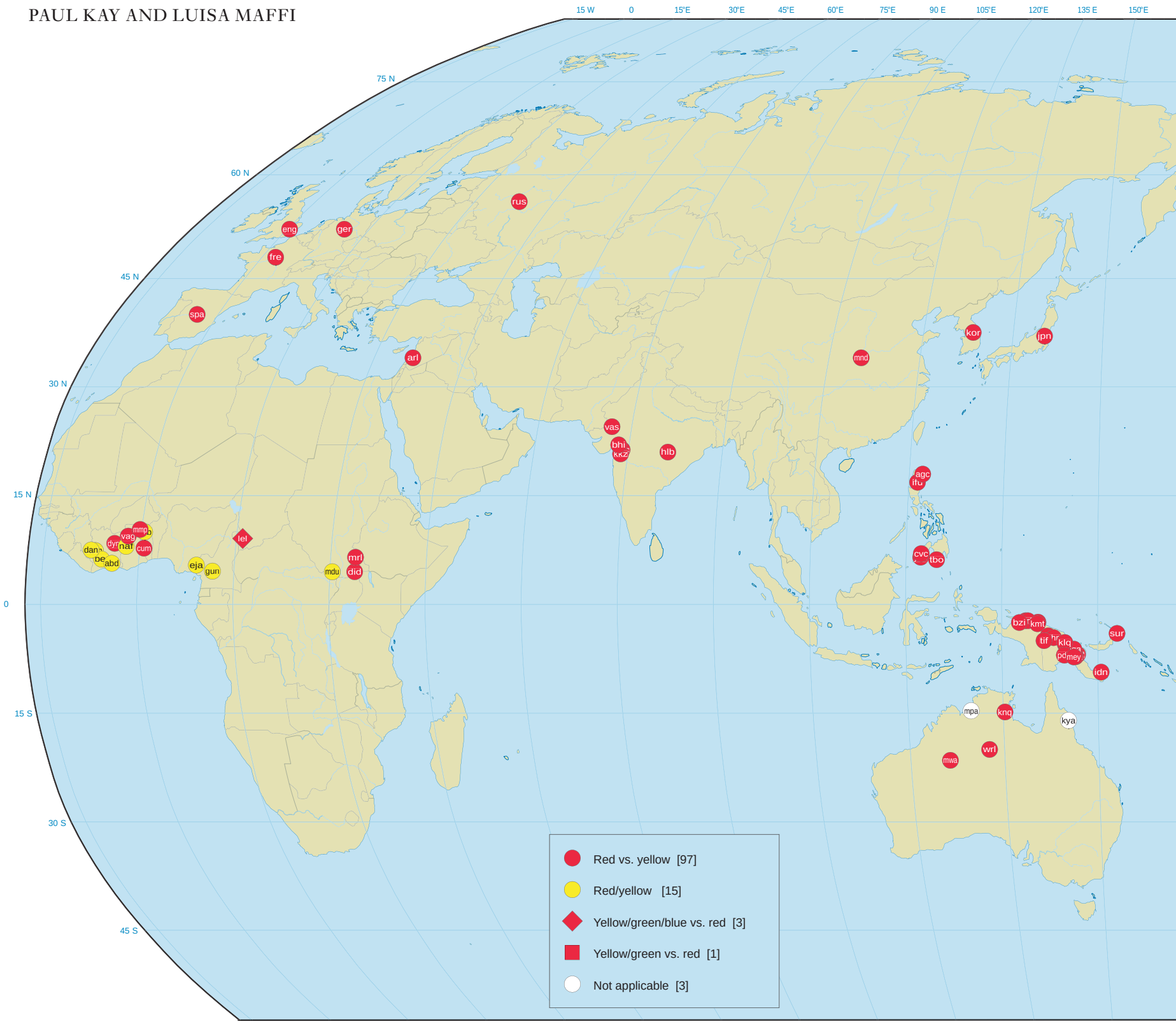
Mesoamerica

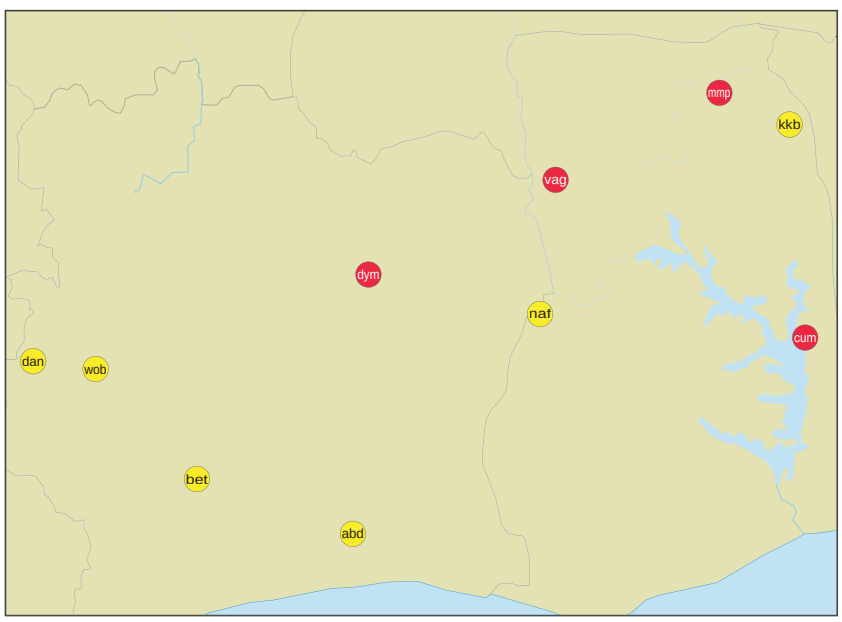


northwestern South America

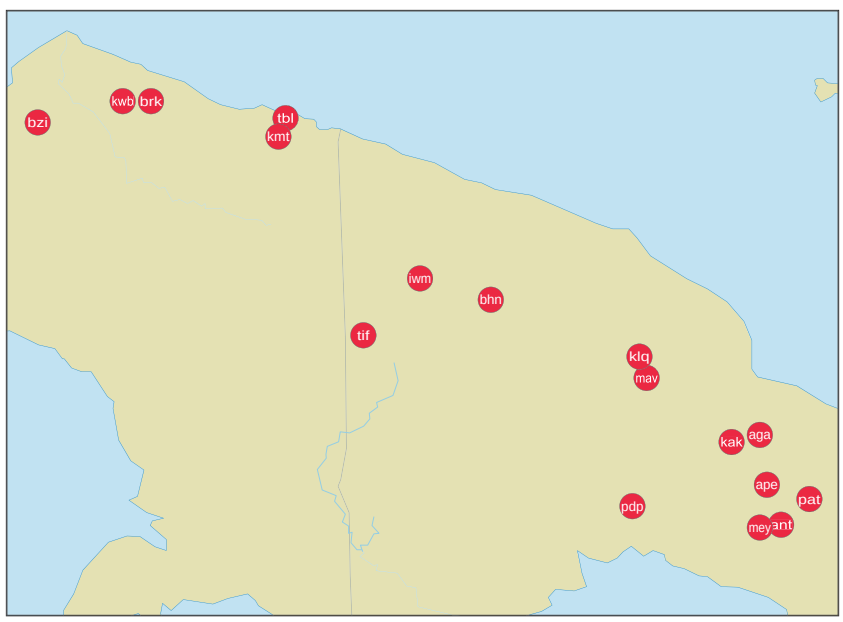
135 Red and Yellow

PAUL KAY AND LUISA MAFFI

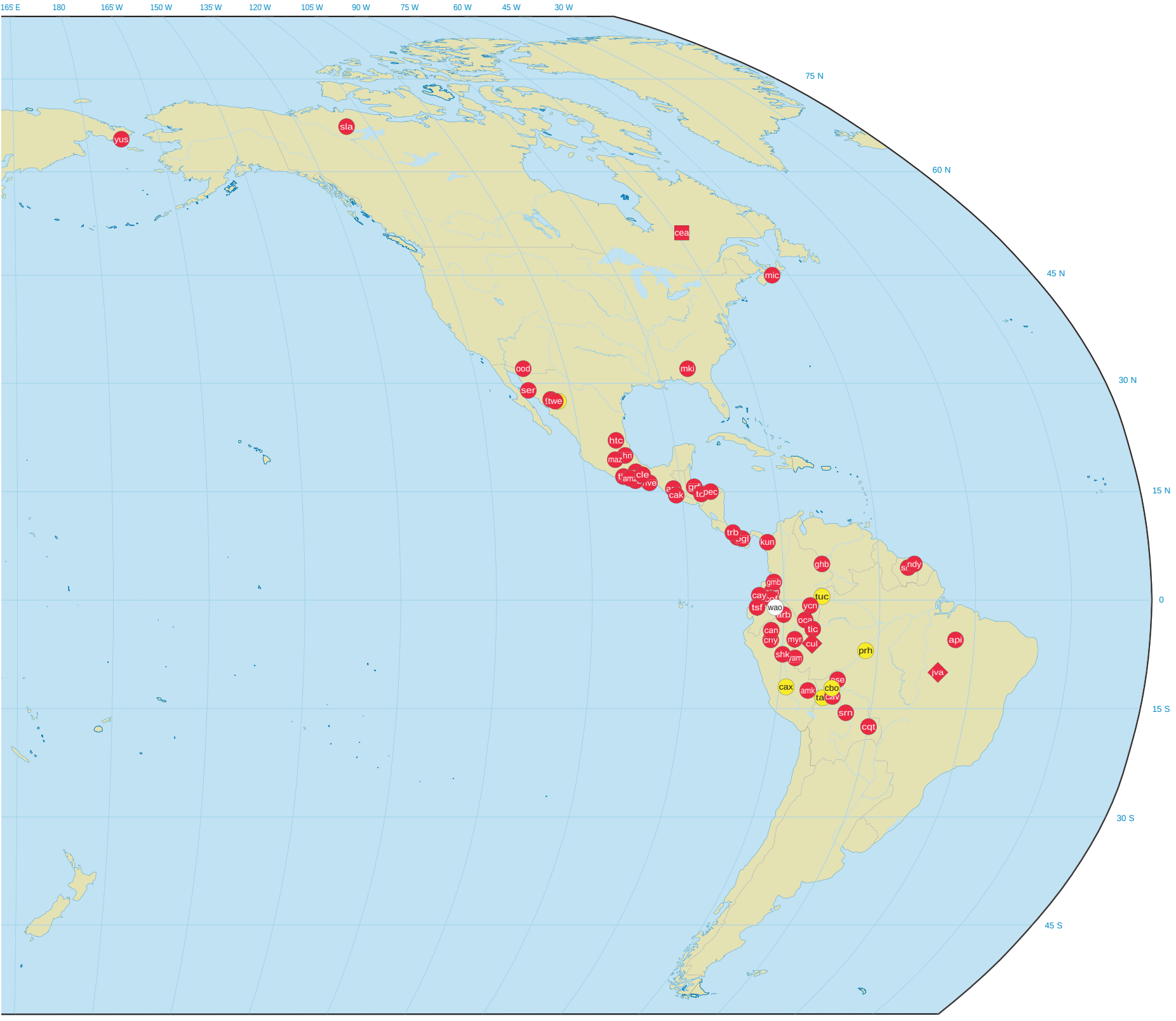


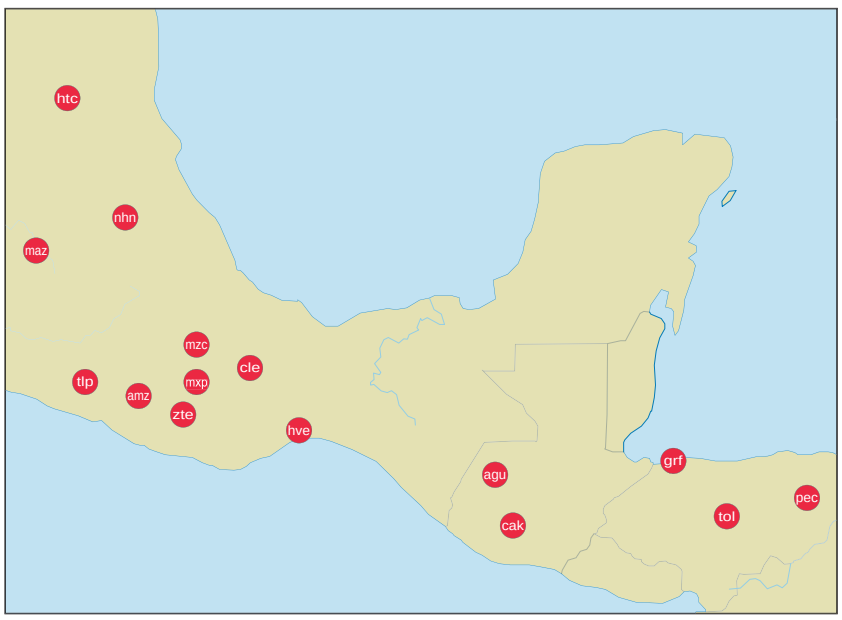


Côte d'Ivoire and Ghana

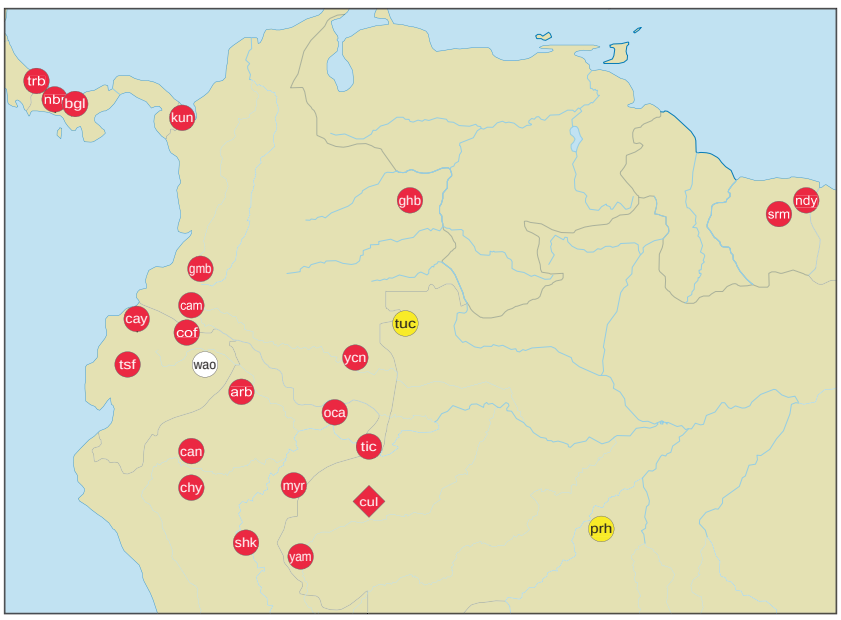


New Guinea





Mesoamerica



northwestern South America

136–137 Personal Pronouns

JOHANNA NICHOLS AND DAVID A. PETERSON

1 Introduction

Personal pronouns, like other closed-class forms, tend to contain nasals and other basic consonants. However, particular paradigmatic combinations of consonants in particular person-number categories are not especially frequent worldwide. There are only two such combinations that show up interestingly on a map: pronoun sets like English *me* and *thee* or Nanai (Tungusic; eastern Siberia) *mi* 'I, me' and *si* 'you', with *m* in the first person and a coronal obstruent in the second, which are common in northern Eurasia and rare elsewhere; and sets like Northeast Maidu (Penutian; California) *ni* 'I, me' and *mi* 'you' with *n* in the first person and *m* in the second, which are common in western America and rare elsewhere. Map 136 here shows the first type, henceforth "*m–T*" patterns, and Map 137 the second, henceforth "*n–m*" patterns.

The maps here are based on the first- and second-person singular forms only. They survey the first consonant in the pronoun root, whether it is word-initial as in Wintu (Penutian; California) *ni* 'I, me; we, us' or follows an initial vowel as in Nez Perce (Sahaptian; Idaho and Oregon) *'iin* 'I, me'. (Whether written or not, the glottal stop that precedes an initial vowel is not considered the first consonant. The Nez Perce form just cited begins with such a glottal stop.) They survey three pronominal forms: **independent pronoun** words like English *I, me, you*; **possessive affixes** as in Hungarian *ház-am* 'my house', *ház-ad* 'your house'; and **verbal affixes** as in Rama (Chibchan; Nicaragua) *n-taak-i* 'I am going', *m-taak-i* 'you are going' (Grinevald n.d.). (Of course, not all languages have all three series.) They consider the two most widely distributed allomorphs of each, e.g. English *I* and *me* for first person singular.

2 Defining "m", "n", and "T"

The consonants mapped are defined as follows:

m = [m] as well as glottalized, voiceless, lenis, and fortis labial nasals (in fact only plain [m] occurred in the language sample)
n = dental or alveolar [n], palatal [ɲ] (and glottalized, voiceless, etc. variants, though none appeared in the sample)
T = any apical obstruent: [t], [d], [s]; palatals such as [č], [š], etc.

Here are some examples of pronouns that fit the two types. First-person *m* and second-person apical obstruent in independent forms:

(1) German
1SG *mich* (accusative)
mein (possessive)
2SG *dich* (accusative)
dein (possessive)

(2) Georgian
1SG *me*
2SG *šen*

(3) Nanai (Tungusic; eastern Siberia)
1SG *mi*
2SG *si*

in possessive affixes:

(4) Hungarian
1SG *ház-am* 'my house'
2SG *ház-ad* 'your house'

in verbal affixes:

(5) Turkish present tense (Lewis 1967: 109)
1SG *alıyor-um* 'I am taking'
2SG *alıyor-sun* 'you are taking'

First person *n* and second person *m* in independent forms:

(6) Wintu (Penutian; California)
1st person *ni*
2nd person *mi*

in possessive affixes:

(7) Pipil (Uto-Aztecan; El Salvador; Campbell 1985: 43)
1SG *nu-chi:l* 'my chili pepper'
2SG *mu-chi:l* 'your chili pepper'

(8) Karok (Hokan; California; Bright 1957: 560)
1SG *nani-tta:t* 'my mother'
2SG *mi-tta:t* 'your mother'

in verbal agreement formatives:

(9) Nisgha (Tsimshianic; British Columbia; Tarpent 1987: 461)
1SG *nə-*
2SG *mə-*
(transitive subject proclitics)

3 Defining the map values

The following values are shown on the maps:

Map 136 M-T Pronouns

○	1. No M-T pronouns	200
●	2. M-T pronouns, paradigmatic	27
○	3. M-T pronouns, nonparadigmatic	3
	total	230

Map 137 N-M Pronouns

○	1. No N-M pronouns	194
●	2. N-M pronouns, paradigmatic	25
○	3. N-M pronouns, nonparadigmatic	11
	total	230

The maps distinguish between **paradigmatic** and **nonparadigmatic** distributions of the consonants. The consonants form a paradigm when they both occur in the same form class(es) of their respective pronouns. For instance, a consistent paradigm occurs in Mapudungun (Araucanian; Chile):

(10) Mapudungun (Smeets 1989)

		1SG	2SG
Independent forms:	Nominative	*iñ̲che*	*ey<u>m</u>i*
	Possessive	*ñ̲i*	*<u>m</u>i*
Verbal suffixes:	Indicative	*-(ü)<u>n</u>*	*-<u>m</u>-i*
	Conditional	*-i-Ø*	*-m-i*

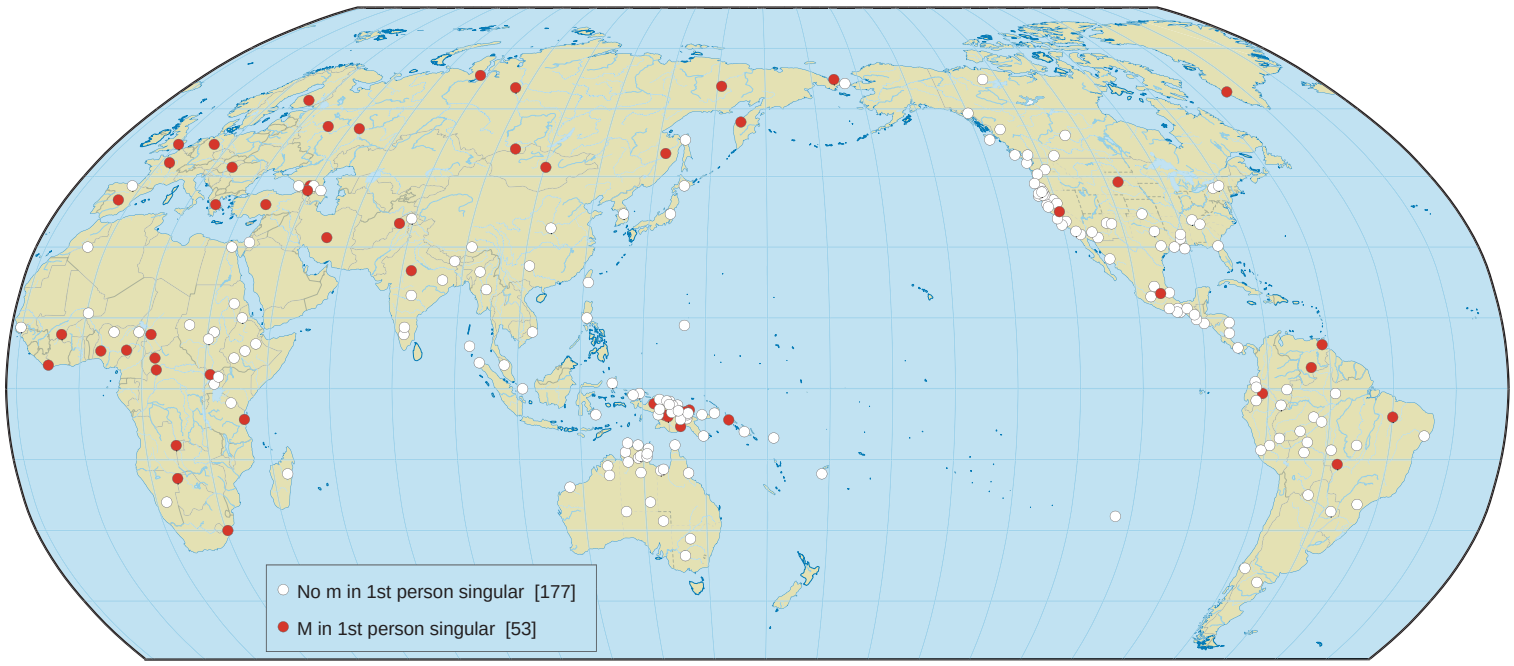
The underlined consonants are in the same position in three different form series and occur in both forms of the set. A subparadigm occurs in Wichí (Mataco-Guaicuru; Argentina):

(11) Wichí (Viñas Urquiza 1974)

		1SG	2SG
Independent Nominative		*oƛam*	*am*
Possessive	animate	*olä*	*alä*
	inanimate	*okä*	*akä*
Verb subject		*o-*	*le-*
Verb object		*-<u>n</u>o*	*-a<u>m</u>*
Verb object		*-o*	*-a*

The first pair of verb-object suffixes (the next to last line here) shows the pattern; no other does. The maps in this chapter do not distinguish consistent paradigms from subparadigms.

A case of first person *n* and second person *m* which do not form a paradigm occurs in Asmat (Asmat-Kamoro; Papua, Indonesia):



Map 136A M in First Person Singular

(12) Asmat (Voorhoeve 1965b)

	1SG	2SG
Independent Nominative	*no-r*	*o-r*
Independent Oblique	*no*	*o*
Verb subject	*-i*	*-em*
Verb object	*-(e)n*	*-(e)n*

Here there is first person *n* in the independent forms and second person *m* in the verb-subject agreement suffix. In fact, *n* occurs in both persons in the verb object forms. There is no place where *n* and *m* co-occur in the same form class.

Finally, the occasional token of one or the other paradigm can turn out to be truly artifactual, as in the following from Grebo (Kru, Niger-Congo; Liberia):

(13) Grebo (Innes 1966)

	1SG	2SG
Emphatic	*mɔ 1*	*mɔ 2/3*
Object	*mo 1/2*	*mo 2/3*
Possessive	*na 1/2*	*na 3*

(numbers indicate tones)

The consonant-vowel forms of the pronouns are the same in all form classes, and only the tones differentiate them. One can find a first person *n* and second person *m* in this set, so this technically counts as a (nonparadigmatic) occurrence of the *n-m* consonantism, but this is rather clearly due to a fluke rather than to a principled linguistic analysis of the material.

It is the paradigmatic sets that are most interesting, and in both maps they are by far the most common of the examples of the two consonant pairs. Nonparadigms (like that of Asmat, above) and flukes (like the forms of Grebo just above) are rare, and occur only outside of the areas where the paradigms are relatively common.

4 Geographical distribution

The pronoun systems in question form two large areal clusters: an *m-T* one centred in northern Eurasia (Map 136) and an *n-m* one extending from western coastal North America to western South America (Map 137). There are a number of other maps in this atlas that have similar areal distributions: the *n-m* concentration on the Pacific Rim and especially the American Pacific Rim recalls the distribution of multiple possessive classification, bound inalienables, numeral classifiers, verb-initial word order, and double marking (see Chapters 59, 58, 55, 81, 23, respectively); the *m-T* cluster in northern Eurasia recalls other Greater Silk Road distributions such as dependent marking, low synthesis, absence of possessive classification, and case-number coexponence (see Chapters 23, 22, 59, 21, respectively).

5 The inset maps

Mapping just *m* in either first-person singular (Map 136A) or second-person singular (Map 137A) yields somewhat more diffuse clusters but reveals the same large areas clearly.

○	1. No *m* in first person singular	177
●	2. *M* in first person singular	53
	total	230

Values of Map 136A M in First Person Singular

○	1. No *m* in second person singular	152
●	2. *M* in second person singular	78
	total	230

Values of Map 137A M in Second Person Singular

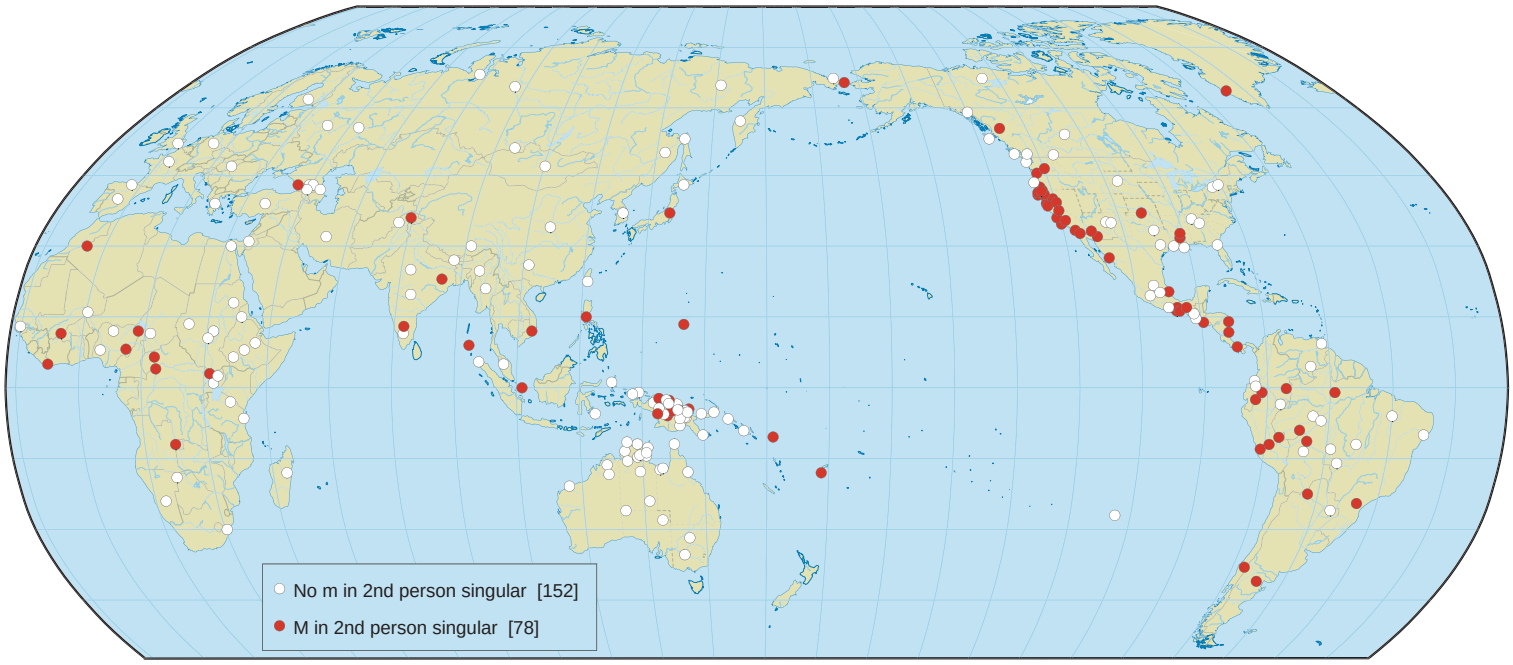
First person *m* (Map 136A) is ubiquitous in the Greater Silk Road area and nearly pan-Eurasian, missing in the Eurasian islands, South-East Asia, and the Asian Pacific Rim; it is well attested in sub-Saharan Africa and New Guinea, sparsely attested in the Americas, and absolutely absent from Australia. Second person *m* (Map 137A) has a solid Pacific Rim distribution with occasional tokens inland in the Americas, in the Eurasian islands, and in Africa; it is absolutely absent from the Greater Silk Road area and Australia.

6 Explanations

What explains these two large unique clusters of languages with the distinctive pronoun consonantisms? The two main explanations that have been offered in the literature are:

(i) Sound-symbolic universals: nasals and other basic consonants are the most elementary consonants, hence the most easily learned by children, and such consonants figure frequently in pronouns and other basic words. While this generalization is correct for broad phonological categories like "nasal", it cannot explain the distributions of the two-consonant paradigms or even of just *m* mapped here, for their distribution is not universal. No explanation based on universals can account for the fact that each of the two paradigms is fairly common in one large area and vanishingly rare elsewhere. The near-absence of each paradigm in every large area but one indicates that they are not due to any underlying tendency but have particular historical origins.

(ii) Genealogical relatedness of the languages: each cluster is made up of the descendants of a single protolanguage. While this explanation may seem plausible, there is no other evidence for the



Map 137A M in Second Person Singular

relatedness of all and only the conforming languages, and the pronoun systems are not sufficient evidence to prove relatedness. The conforming languages in each area belong to several different well-established families:

m–T:

Language(s)	Family
Eurasia:	
Chukchi, Itelmen	Chukotko-Kamchatkan
Chuvash, Turkish, Tuvan	Turkic
Finnish, Hungarian, Nenets	Uralic
French, German, Modern Greek, Hindi, Ossetic, Persian, Russian, Spanish, Waigali	Indo-European
Georgian	Kartvelian
Khalkha	Mongolic
Nanai	Tungusic
Kolyma Yukaghir	(isolate)
Elsewhere:	
Nigerian Fula, Grebo	Niger-Congo
Lakhota	Siouan
Salt-Yui	Chimbu
Southern Sierra Miwok	Utian
Usan	Adelbert Range

n–m:

Language(s)	Family
Americas:	
Axininca Campa	Arawakan
Cashinahua	Panoan
Chimariko	(isolate)
Upper Chinook	Chinookan
Highland Chontal	Tequistlatecan
Mesa Grande Diegueño	Yuman
Kiowa	Kiowa-Tanoan
Luiseño, O'odham, Pipil, Yaqui	Uto-Aztecan
Northeast Maidu	Maiduan
Mapudungun	Araucanian
Miskito	Misumalpan
Mixe	Mixe-Zoque
Rama	Chibchan
Northern Sahaptin	Sahaptian
Wari'	Chapacura-Wanhan
Wichí	Matacoan
Wintu	Wintuan
Yawelmani	Penutian
Elsewhere:	
Dumo	Sko
Supyire	Niger-Congo

On the other hand, in each cluster there are a few language families that are likely to be deeply related. Indo-European and Uralic are likely to be related, though the relatedness is so distant that we may never have sufficiently firm evidence to prove it by conventional means (e.g. Ringe 1998, Oswalt 1991). Still deeper relatedness of several of the other *m–T* families of Eurasia to these has been proposed but never demonstrated. In any event the first person *m* is secondary in Turkic and Tungusic (where it developed from a *b*). In North America, a set of language families which have the *n–m* system is thought to be possibly deeply related: Utian, Wintun, Maidun, Sahaptian, and possibly also Chinookan may belong to a proposed megafamily called Penutian. But not all of the putative Penutian daughter families exhibit the system, and the putative Penutian set represents less than half of the tokens of *n–m* systems. Furthermore, one putative Penutian language, Southern Sierra Miwok, exhibits the *m–T* system. Thus deep genetic relatedness is no explanation for the pronoun patterns.

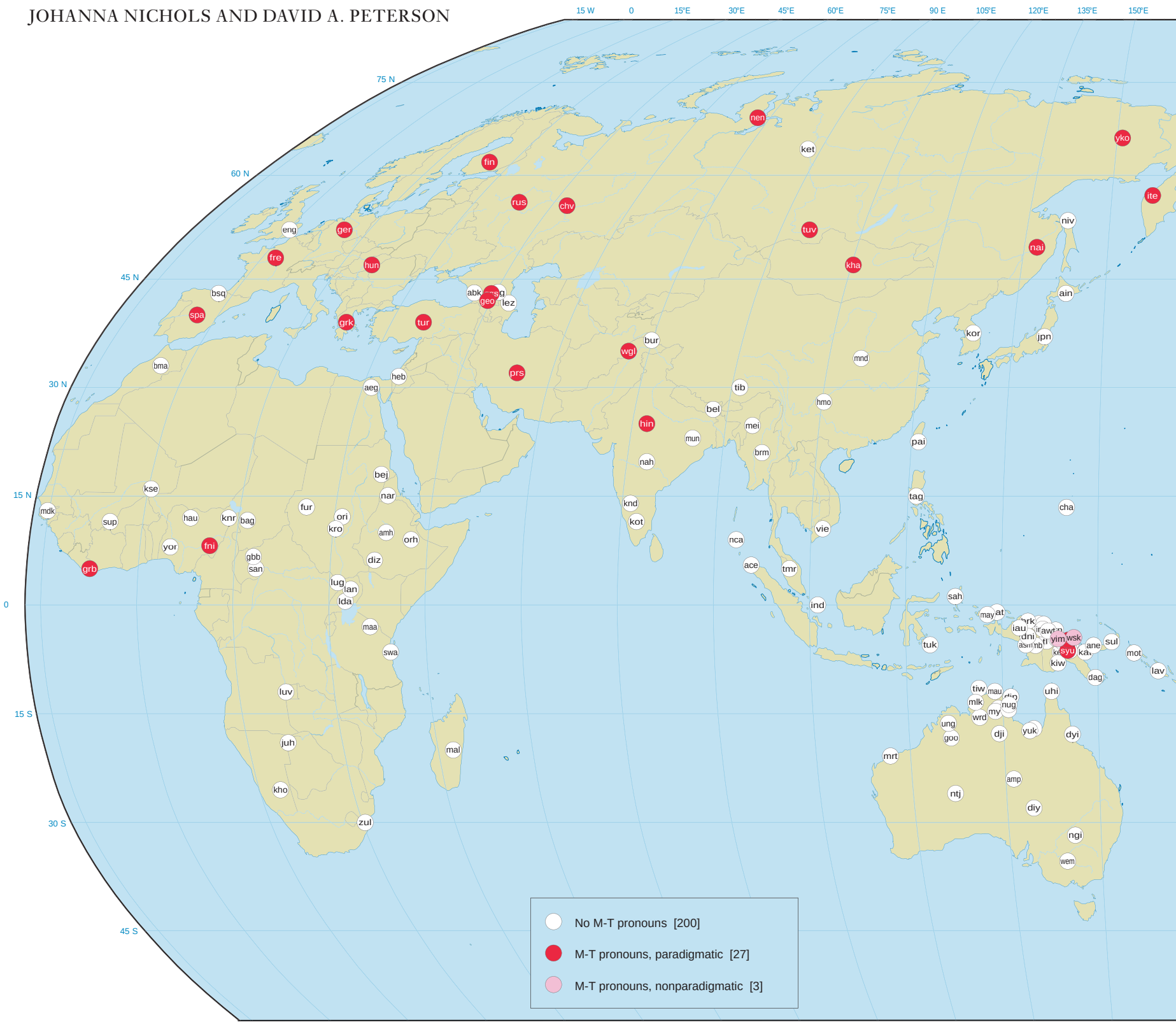
Both the *m–T* and the *n–m* paradigms are found in areas where spreads are known to have been centred—the Greater Silk Road and the Pacific Rim—and both appear to be the products of geographical spread rather than just universals or just inheritance. Such spreads of pronominal consonantism patterns evidently do not occur often: we have only these two clear cases in the whole world. The origins of both are old (older than any of the individual families exhibiting the systems, for instance): the Eurasian *m–T* pattern is at least pre-Neolithic (Nichols 2001), and the American *n–m* one is late glacial at the latest (Nichols and Peterson 1996). Since these ages are greater than the time depth to which linguistics can generally trace genealogical relations, the details of their origins are lost in time.

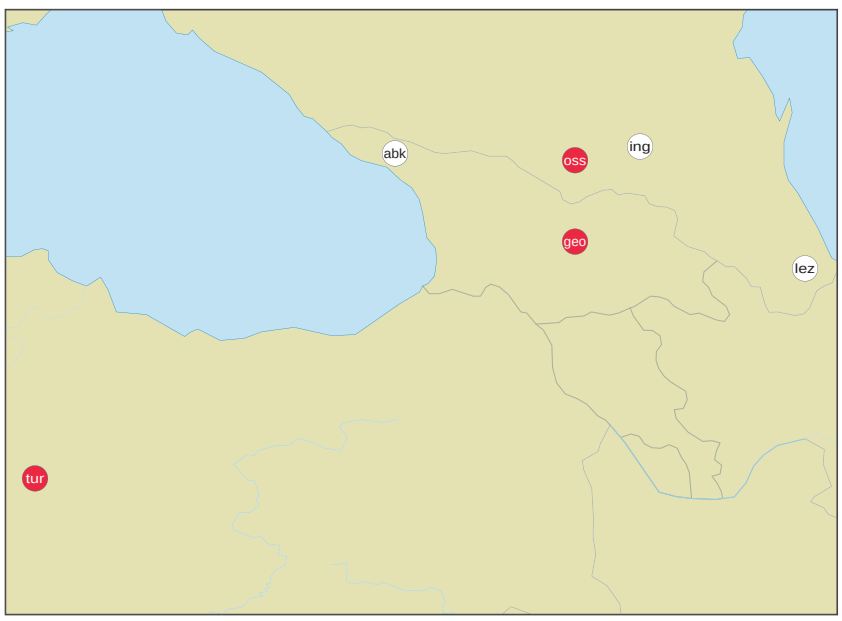
We also looked for other interesting consonant patterns in pronominal roots, e.g. those involving velar nasals and velar obstruents *k, g*, etc. However, no such patterns showed conspicuous large areal distributions; all are more or less evenly distributed worldwide. So far, then, the *m–T* and *n–m* paradigms mapped here are unto themselves.

For more discussion see Nichols and Peterson 1996 and references therein; for subsequent discussion, see Nichols 2001.

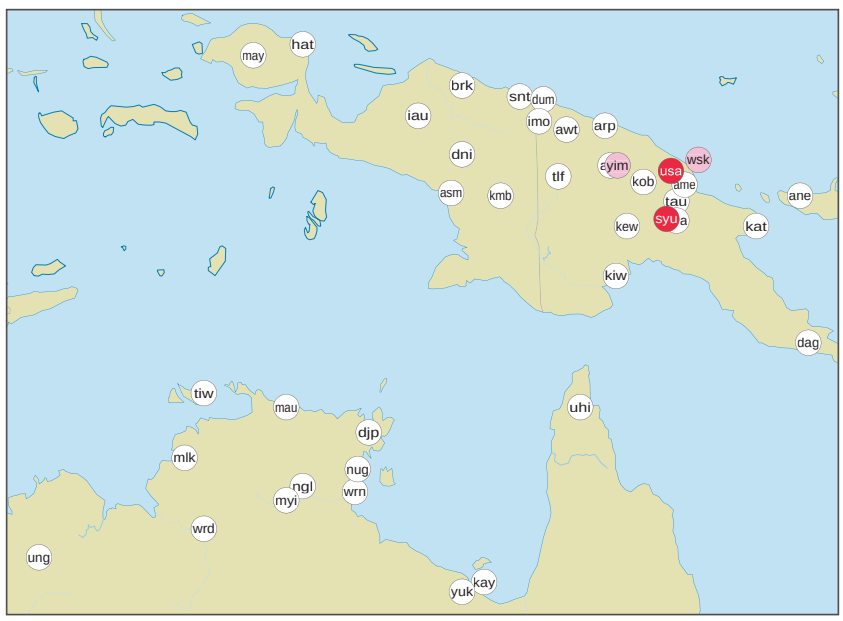
136 M-T Pronouns

JOHANNA NICHOLS AND DAVID A. PETERSON

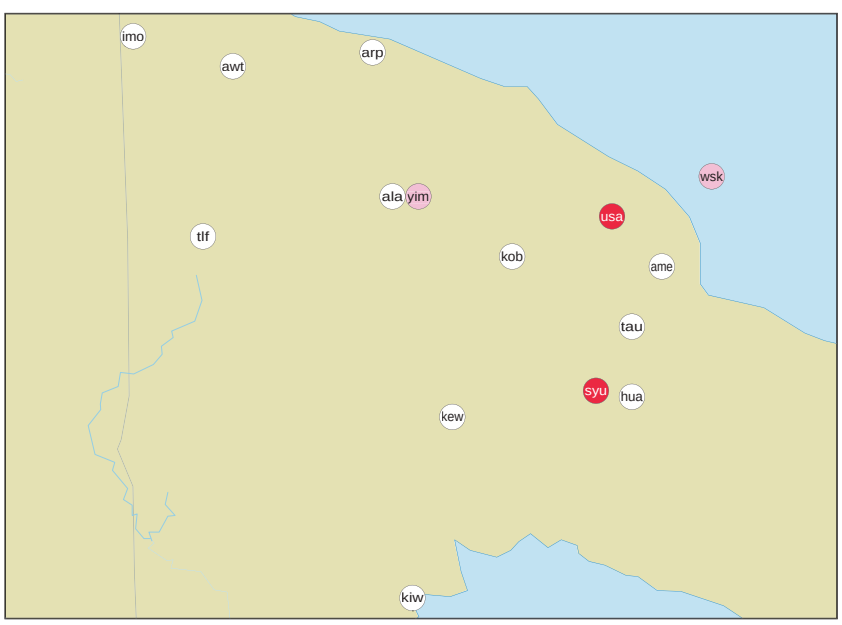




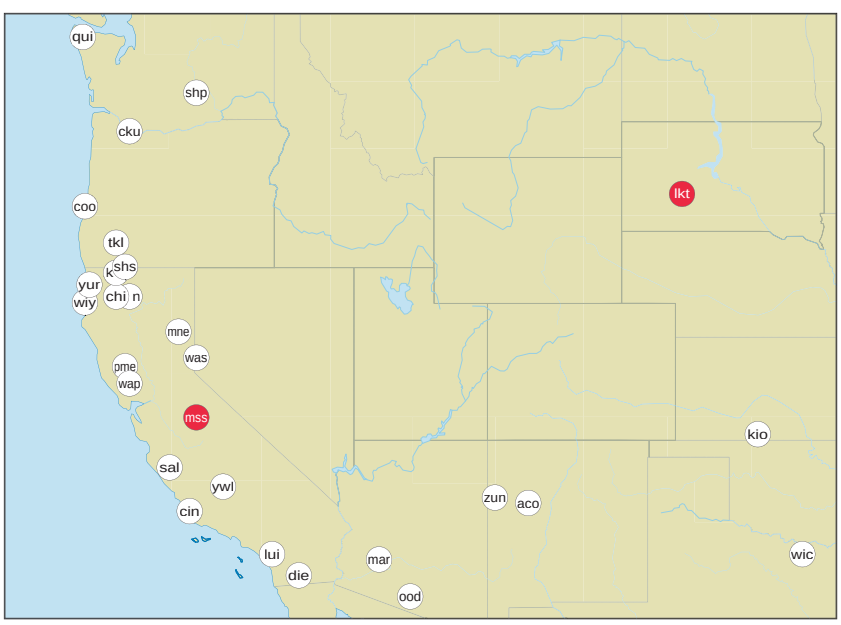
Caucasus and adjacent areas



New Guinea and northern Australia



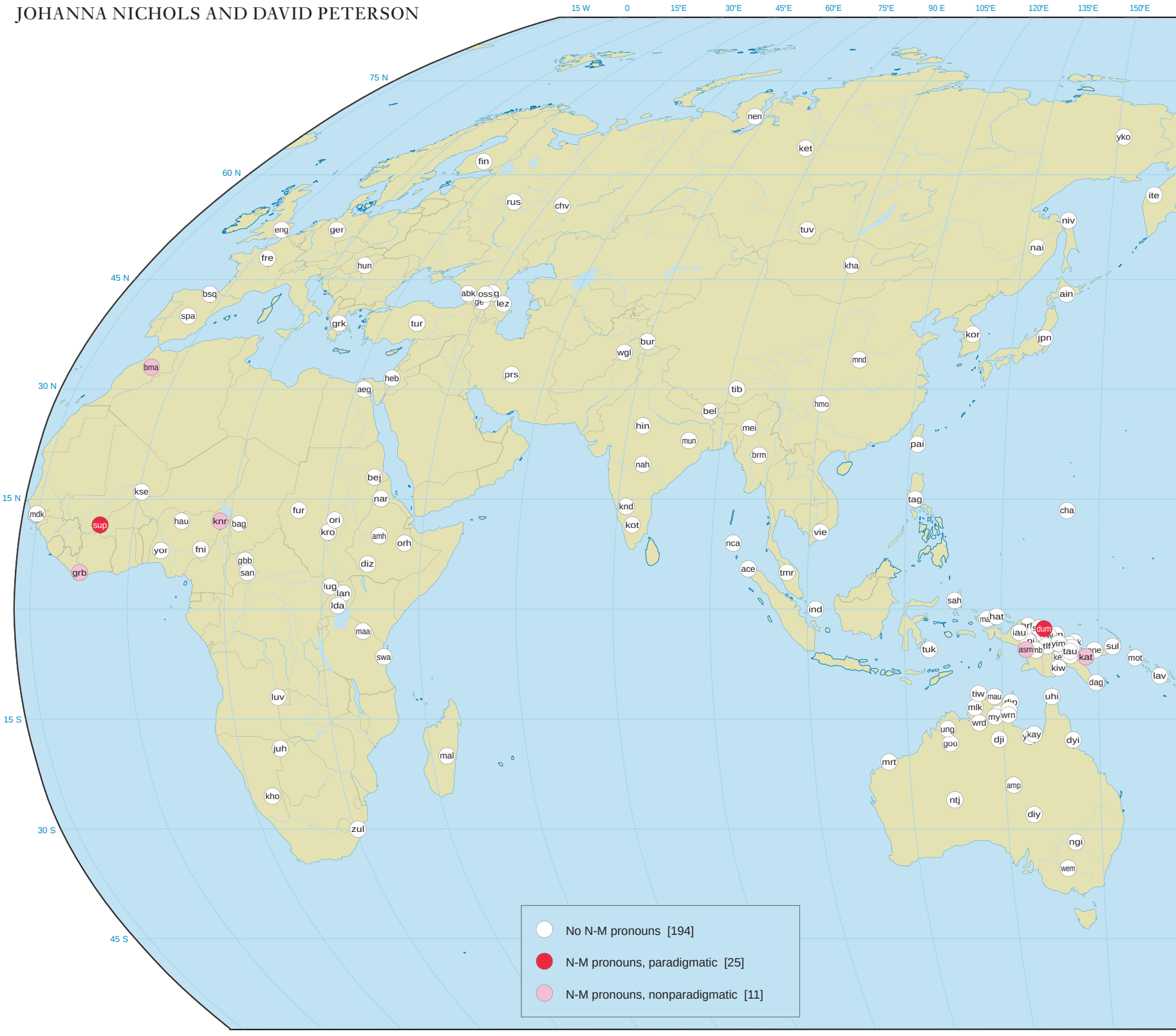
central New Guinea

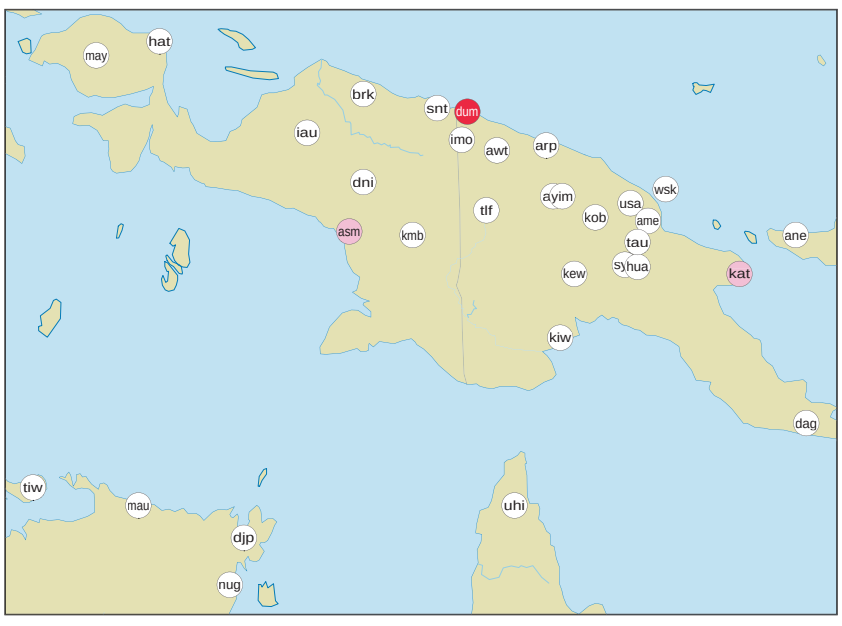


western United States

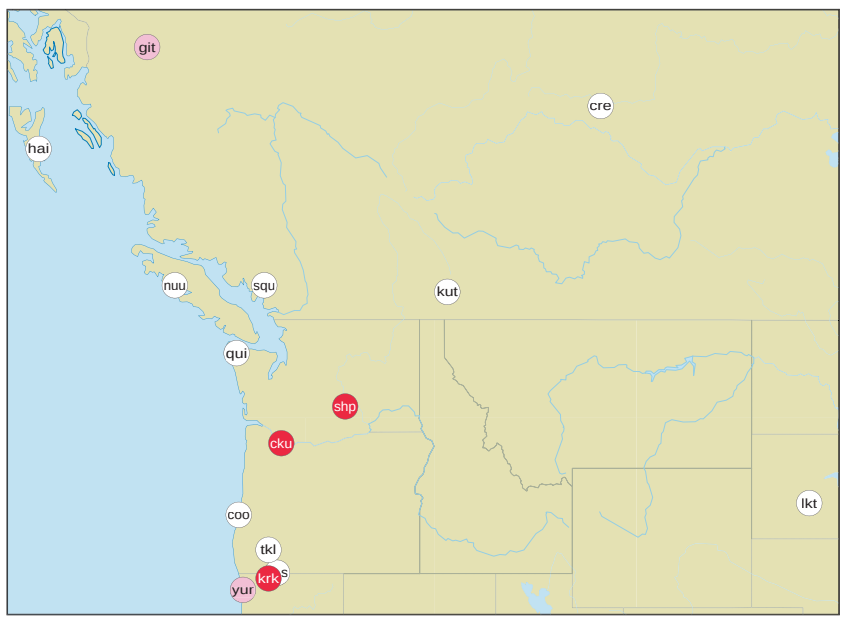
137 N-M Pronouns

JOHANNA NICHOLS AND DAVID PETERSON

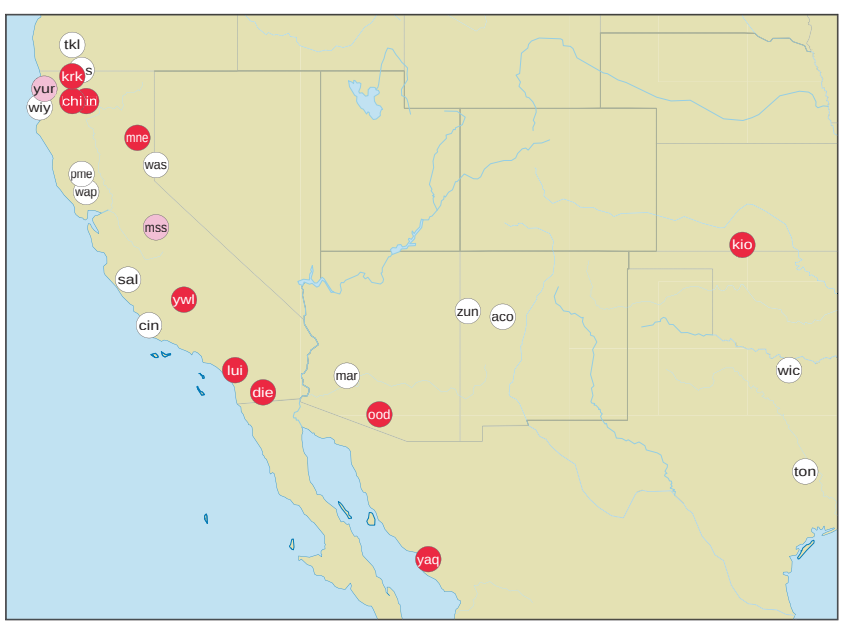




New Guinea and northern Australia



western Canada and United States



southwestern United States and northern Mexico

southern Mexico and Central America

138 Tea

ÖSTEN DAHL

This chapter differs from most of the chapters in the atlas in several ways, the focus being on the actual phonological shape of a lexical item expressing a particular concept—moreover, one that is definitely not part of the universal heritage of mankind but rather is connected to a specific agricultural product. What is shown on the map are the sources of the words for 'tea' in various languages in the world, illustrating in a somewhat peculiar way the spread of words together with material culture. In particular, the map shows the importance of long-distance contacts, such as the trade relations between European countries and East Asia, and is thus a healthy reminder of the fact that languages need not be geographically contiguous to influence each other. Also, in contrast to, for example, the patterns illustrated in the chapters on M-T and N-M pronouns (Chapters 136–137), which presumably go far back in linguistic prehistory, the distribution of words for 'tea' depends on quite recent historical processes.

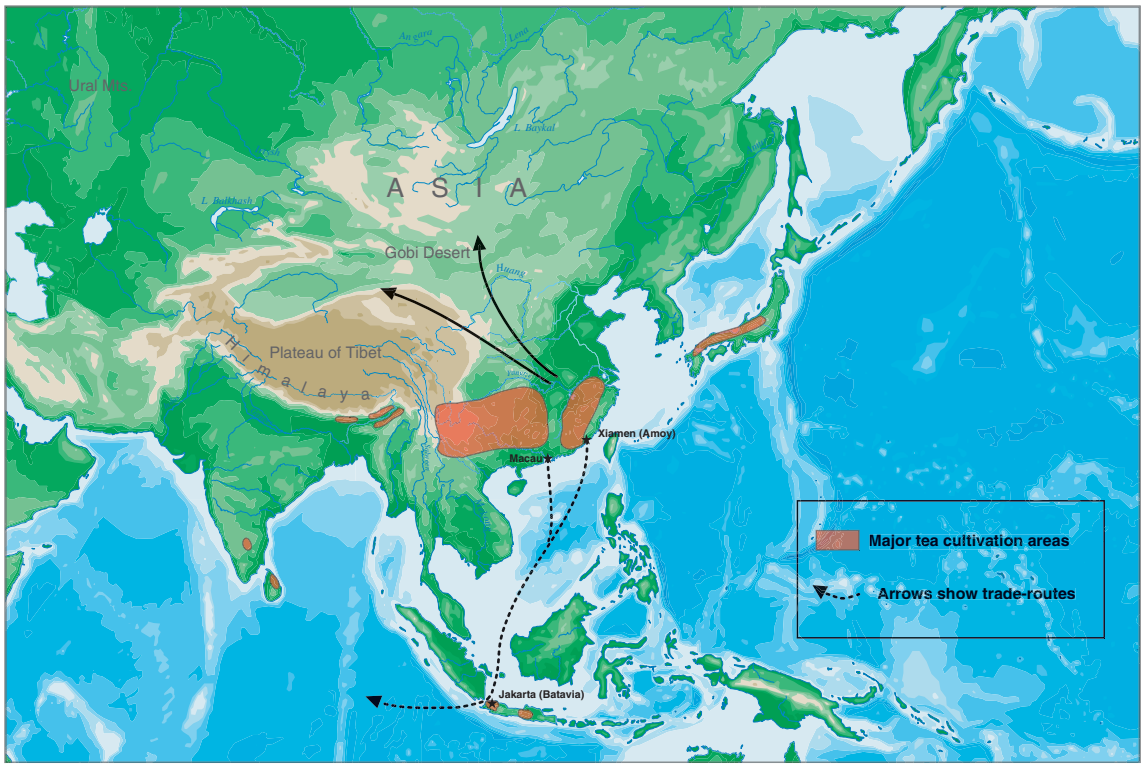
1.	Words derived from Sinitic *cha*	109
2.	Words derived from Min Nan Chinese *te55*	81
3.	Others	40
	total	230

The tea plant (*Camellia sinensis*) grows in a vast area that stretches from Assam (India) in the west to the east coast of China and southwards into Myanmar, Thailand, and Vietnam. The use of the beverage made from the leaves is said to have originated in China and according to legend goes back to 2700 BCE. Tea spread first to Japan, around 800 CE. In the seventeenth century the use of tea became common in Europe.

Most words for 'tea' found in the world's languages are ultimately of Chinese origin, but they differ significantly in their form due to their coming via different routes. The differences begin already on Chinese soil. Most Sinitic languages have a form similar to Mandarin *chá*, but Min Nan Chinese, spoken e.g. in Fujian and Taiwan, has instead forms like *te55* (Chaozhou). The Dutch traders, who were the main importers of tea into Europe, happened to have their main contacts in Amoy (Xiamen) in Fujian. This is why they adopted the word for 'tea' as *thee*, and in this form it then spread to large parts of Europe. The influence from Amoy is also visible in many languages spoken in the former Dutch colonies, as in Malay/Indonesian and Javanese *teh*. However, the first European tea importers were not the Dutch but the Portuguese, in the sixteenth century; their trade route went via Macao rather than via Amoy, and consequently Portuguese uses *chá*, derived from Cantonese *cha*. Likewise, languages spoken in those countries in eastern Europe and Asia which got their tea overland rather than from the Dutch tend to use forms such as *chai*. In Standard English, the [e:] vowel changed into [i:] (although a form conventionally spelled *tay* is preserved in some dialects) by a general sound change, and this pronunciation is reflected in many languages that took over the word from English, e.g. Yoruba (Nigeria) *tii*, !Xóõ (Botswana) *tîi*, Cocopa (California and northern Mexico) *ti*.

In spite of the general tendency for words for 'tea' to be borrowed, a significant number of languages use native denominations, particularly within the area where the tea plant grows naturally. Beverages made by infusion from leaves of various plants are of course common in many places, and some languages may have extended the use of a word for these to the imported product. Conversely, words originally used for tea have been extended to other similar drinks, which creates some difficulty in consulting dictionaries, since it is not always clear if a word denotes specifically tea or some other beverage.

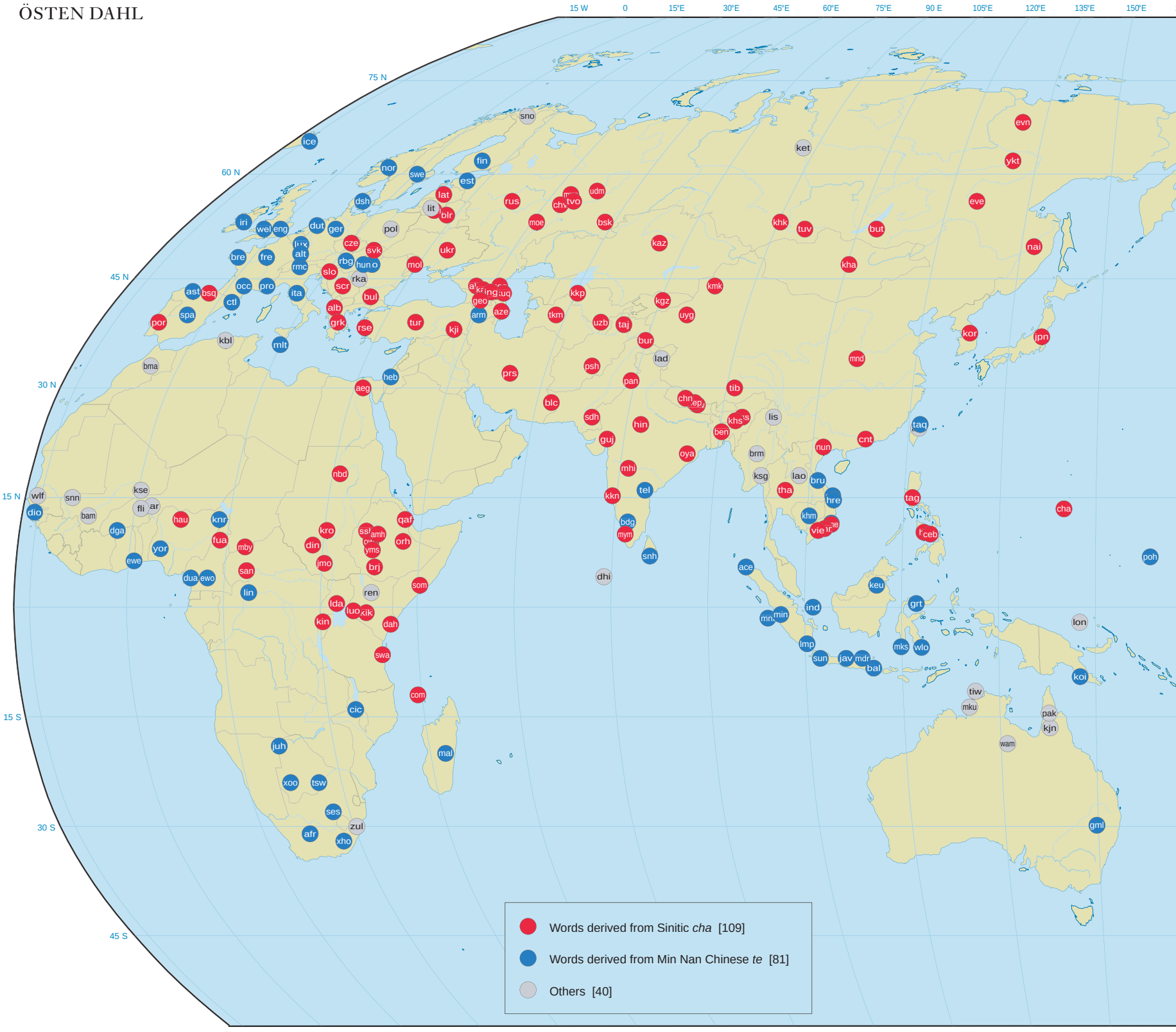
Many dictionaries, especially of languages spoken outside the major tea-consuming countries, do not list any word for 'tea'. It is in general impossible to determine whether this results from the dictionary being incomplete or from there being no word for 'tea' in the language. Probably the situation is often an intermediate one: tea is only marginal in the culture, and whenever there is need to refer to it, a word from the dominant language is used.

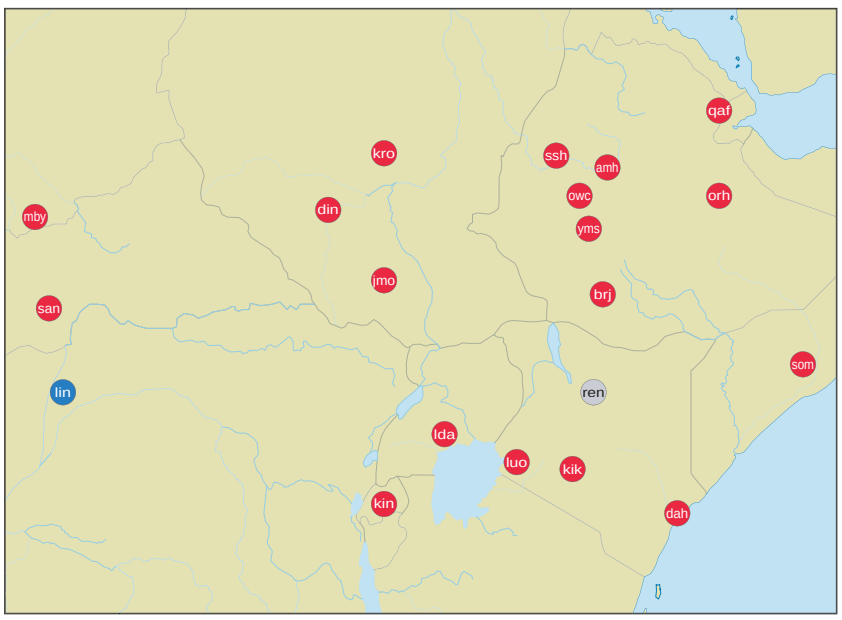


Map 138A Major Tea Cultivation Areas and Trade Routes

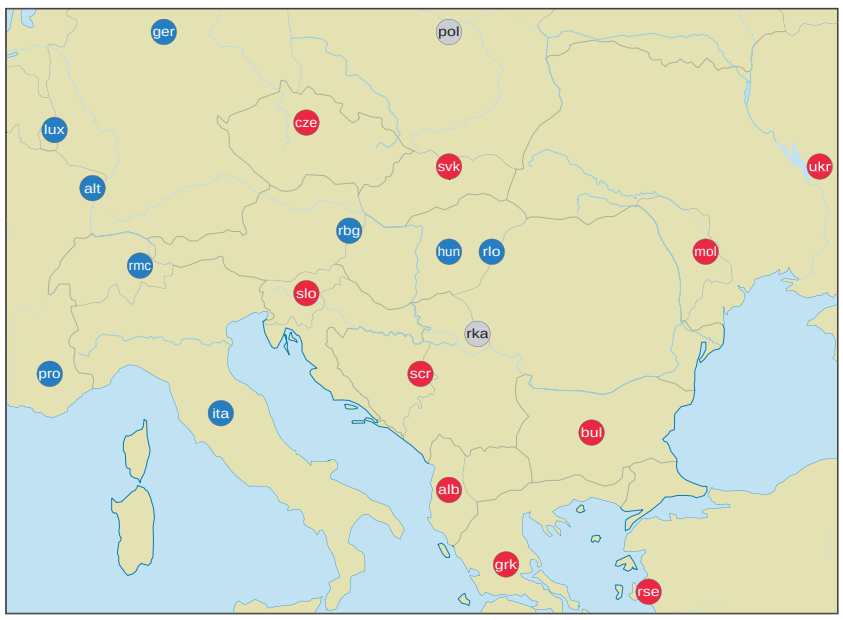
138 Tea

ÖSTEN DAHL

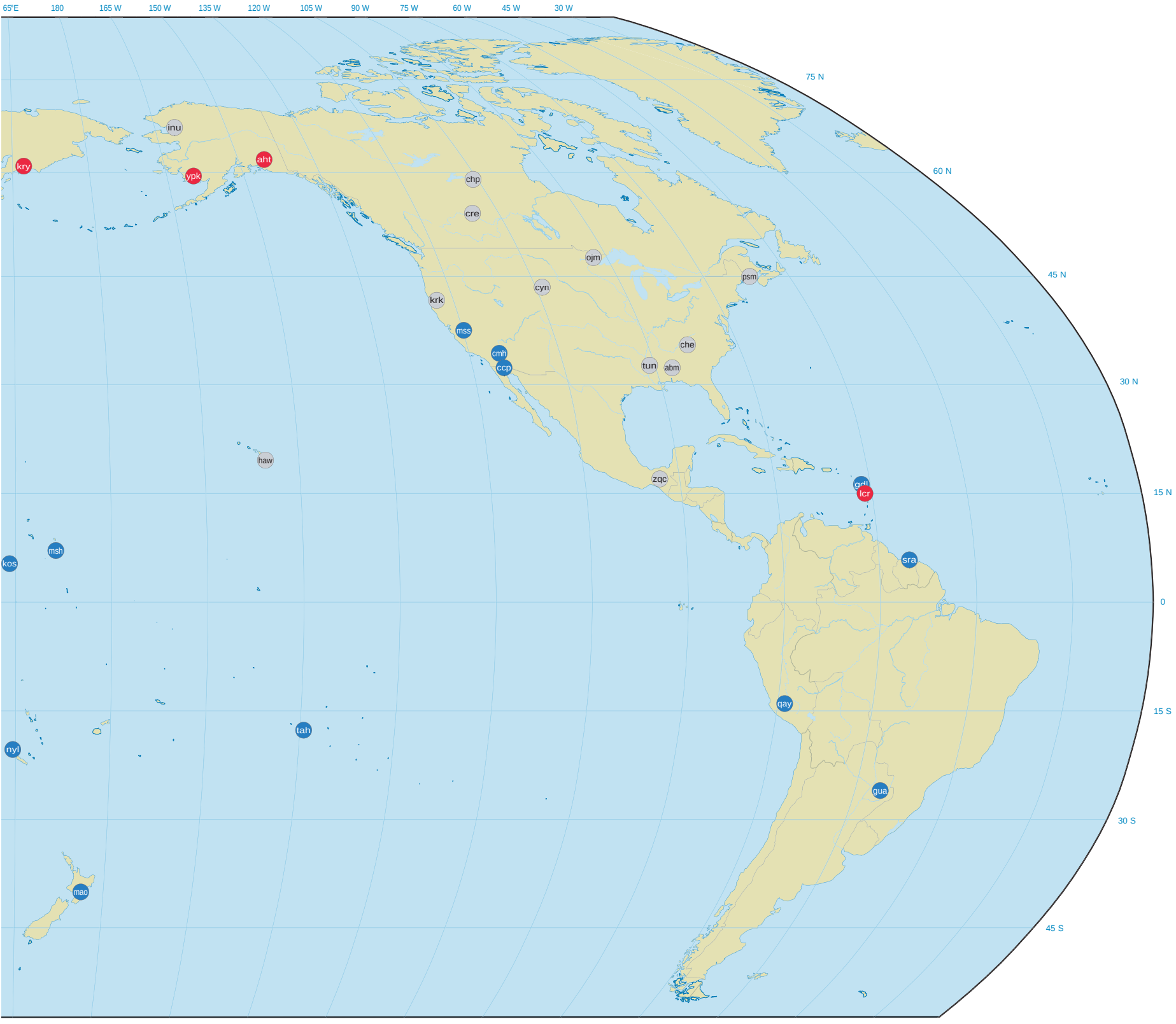


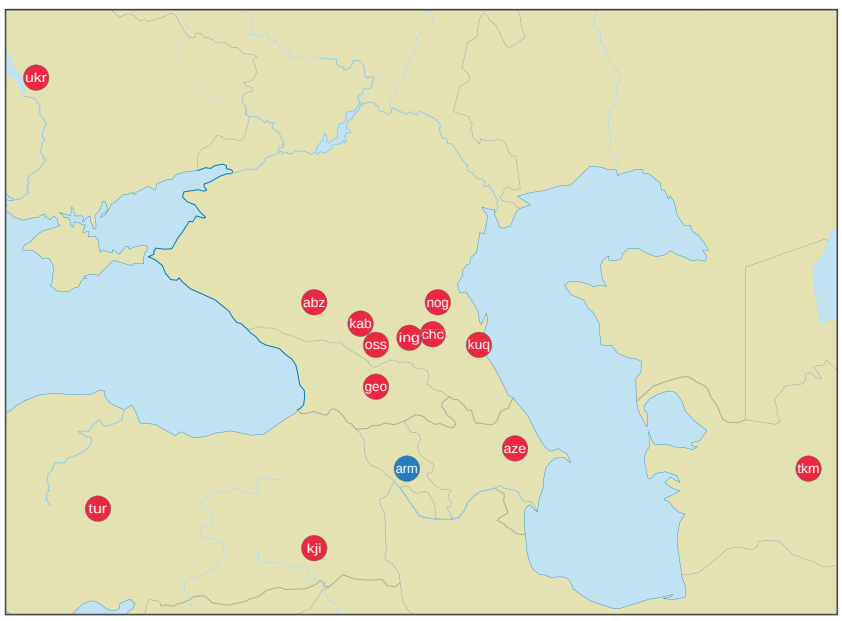


eastern Africa

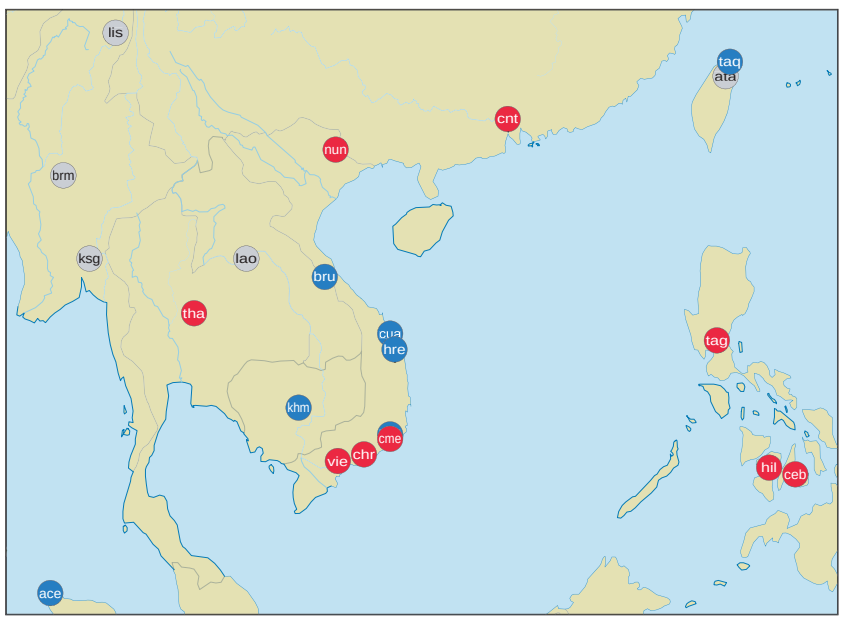


southeastern Europe





Caucasus and surrounding areas



northern Southeast Asia

Sign Languages

ULRIKE ZESHAN

1 Sign languages

Sign languages are natural human languages that have arisen wherever deaf people have come together in sufficient numbers to form a linguistic community. From the 1960s onwards, sign-language researchers have demonstrated that sign languages are fully complex human languages with an intricate grammatical organization of their own and are in every way to be considered on a par with spoken languages (e.g. Klima and Bellugi 1979; Boyes Braem 1990; Sutton-Spence and Woll 1999). They do, however, constitute a separate linguistic type of their own by virtue of the fact that information is transmitted in a visual-gestural modality rather than a vocal-auditory modality.

Although sign-language research has made much progress in the past decades, most sign languages in the world remain hardly described or entirely undocumented. The bulk of linguistic work has been done on only a few sign languages in industrialized countries, in particular in the United States, while barely anything is known about most sign languages in Asia, Africa, South America, and Central America.

2 Sign-language typology

Although sign languages are of great potential interest to linguistic typology, they have so far not figured in any cross-linguistic typological survey, chiefly due to the unavailability of data. The two chapters on sign languages in this volume have resulted from the first large-scale typological survey of language structures across sign languages around the world, investigating negative and interrogative constructions (cf. Zeshan 2004a, 2004b). The chapters on question particles and irregular negatives are based on these data, which have mostly been specifically generated for the survey and have previously been unpublished.

Since it has only just become possible to do typological surveys of sign languages, the criteria for dealing with the data are necessarily different from the spoken language surveys in this volume. Rather than sampling the languages, as is customary for spoken language typology, any and all available information has been included in the sign language survey. This amounts to the 39 sign languages that are listed in Table 1 below.

Most sign languages in the data are used as primary languages in urban deaf communities, and §§3 and 4 below apply to this situation only. In addition, there are two contact sign varieties and three village-based sign languages in the data. Plains Indians Sign Language, now an endangered language, is a contact sign variety used among a number of North-American Indian tribes whose spoken languages are mutually unintelligible (Taylor 1996). International Sign is a sign pidgin that has developed in international settings as a result of contact between deaf users of different sign languages from Europe and North America (Webb and Supalla 1994). Urubú Sign Language, used among an Indian tribe in Brazil (first reported by Kakumasu 1966), Kata Kolok in Bali (Branson et al. 1996), and Adamorobe Sign Language in Ghana (Frishberg 1987) are village-based sign languages that have arisen locally due to a high proportion of deaf individuals.

Although the availability of data necessarily results in an overrepresentation of European sign languages, all regions of the world are represented in the data. The regional distribution of sign languages amounts to fifteen European sign languages, eleven Asian sign languages, six sign languages from the Americas, four African sign languages, and two Australasian sign languages. Relationships between the sign languages in terms of genealogical grouping and language contact are described in the next section.

Table 1 Sign languages in the typological survey

Sign language	*Country/region*
Adamorobe Sign Language	Adamorobe village (Ghana)
American Sign Language (ASL)	United States, Canada except Québec
Auslan	Australia
British Sign Language (BSL)	Great Britain
Chinese Sign Language	mainland China
Deutsche Gebärdensprache (DGS)	Germany
Finnish Sign Language (Suomalainen viittomakieli)	Finland
Greek Sign Language	Greece
Hong Kong Sign Language	Hong Kong (China)
Indo-Pakistani Sign Language	India, Pakistan
International Sign	
Irish Sign Language	Ireland
Islenskt Taknmal (Icelandic Sign Language)	Iceland
Israeli Sign Language	Israel
Kata Kolok	Desa Kolok village (Bali)
Kenyan Sign Language	Kenya
Langue des Signes Française (LSF)	France
Langue des Signes Québecoise (LSQ)	Québec (Canada)
Lengua de Señas Argentina	Argentina
Lengua de Señas Española	Spain except Catalonia
Língua de Sinais Brasileira	Brazil
Língua Gestual Portuguesa	Portugal
Lingua Italiana dei Segni (LIS)	Italy
Lughat al-Isharat al-Lubnaniya	Lebanon
Nederlandse Gebarentaal	Netherlands
New Zealand Sign Language (NZSL)	New Zealand
Nihon Shuwa (Japanese Sign Language)	Japan
Norsk Tegnspråk	Norway
Plains Indians Sign Language	United States
Russian Sign Language	central part of Russia
South Korean Sign Language	South Korea
Svenska Teckenspråket	Sweden
Taiwanese Sign Language (Ziran Shouyu)	Taiwan
Tanzania Sign Language (Lugha ya Alama Tanzania)	Tanzania
Thai Sign Language	Thailand
Türk İşaret Dili	Turkey
Ugandan Sign Language	Uganda
Urubú Sign Language	Brazil
Vlaamse Gebarentaal	Flemish part of Belgium

3 Relationships between sign languages

Relationships between languages are of two types: genealogical relationships and language contact. The first type of relationship pertains to language families, while the second type of relationship has to do with areal typology. However, the kinds of situations known to hold between spoken languages do not necessarily present themselves in the same way for sign languages. It is not at all clear how genealogical relationships should be defined and described for sign languages, and no one has attempted to apply the familiar historical-comparative method to sign languages. In fact, this may well be impossible for a number of reasons. For example, no equivalent to regular sound change has yet been found in any sign language.

It is, however, possible to link sign languages to each other because, as far as we know, most sign languages in their present form are young languages with a documented history. This history is often linked to, but not limited to, the beginnings of special education systems for the deaf. In a number of cases, special education for the deaf was first brought to a country from abroad, including the importation of some influence from the sign language of that other country. This potentially results in creolization of pre-existing

indigenous forms of sign languages, sign pidgins, or so-called "home sign" systems (the latter two being less developed forms of gestural communication) with a foreign sign language. For example, American Sign Language is believed to have arisen from a creolization situation involving French Sign Language and pre-existing local sign varieties. While relationships between sign languages can thus be posited on the basis of historical knowledge, it is not clear whether these relationships can be considered "genealogical" in the same sense of the term as it is applied to spoken languages. With all due caution, associations among some of the sign languages in the data are proposed in Table 2, without attempting a more sophisticated definition of the kind of relationship involved.

Table 2 Relationships between sign languages

British Sign Language—Auslan—New Zealand Sign Language
Nihon Shuwa—Taiwanese Sign Language—South Korean Sign Language
Langue de Signes Française—American Sign Language—Russian Sign Language—Nederlandse Gebarentaal (Northern dialect)—Vlaamse Gebarentaal—Langue des Signes Québecoise (LSF and ASL influence)—Irish Sign Language—Língua de Sinais Brasileira
Deutsche Gebärdensprache—(perhaps also other sign languages in Europe and the Middle East)—Israeli Sign Language
Svenska Teckenspråket—Finnish Sign Language
American Sign Language—Ugandan Sign Language—Thai Sign Language—Kenyan Sign Language
Lingua Italiana dei Segni—Lengua de Señas Argentina
Chinese Sign Language—Hong Kong Sign Language

As a general caveat to Table 2, it should be noted that the evidence is often more of an anecdotal nature rather than being the result of specific linguistic or historical research, which has yet to be undertaken in almost all cases. Where historical connections are not known or have yet to be discovered, there is currently no established way of positing relationships between sign languages. And it is of course entirely possible for a full-fledged sign language to arise on its own, without any influence from elsewhere.

The effects of language contact between sign languages are also barely understood at present. Sign languages seem to be very dynamic, and readily share and borrow both lexical and grammatical material almost on a worldwide basis. It seems to take very little time for a sign language to change considerably, and the impact of language contact can be quite noticeable over a relatively short period of time. Thus American Sign Language has had considerable influence on a number of sign languages worldwide during the period after the Second World War. Language contact between signed and spoken languages is described in the following section.

4 Relationships between signed and spoken languages

Since sign languages in any country constitute a linguistic minority, practically all sign-language users are bilingual to some extent in the sign language and the surrounding spoken language or languages. Another link between the signed and the spoken language is constituted by the hearing children of deaf parents, who grow up using both sign and speech and are often fully bilingual in both languages. Deaf people's spoken language competence and hearing people's sign language competence, where the sign language is acquired as a second language, vary considerably and are often rather poor, especially in countries with few educational and technical resources. Nevertheless, the constant contact with the spoken language always opens up the possibility of spoken-language influence on the sign language and borrowing of linguistic material from the spoken language. The mechanisms, extent and history of this kind of language contact are, however, far from being well understood for most sign languages in the world.

For the purposes of the sign language chapters in this section, we need to focus on a particular type of language contact, the use of contrived sign systems intended to mirror the structure of a spoken language. This is certainly one of the most powerful ways in which a spoken language can exert influence on a sign language, but since such systems are not in use everywhere, their existence also involves interesting typological questions. Contrived hybrid sign systems are predominantly used in educational settings, often as a tool intended to teach the structure of the spoken language. The systems are known by names such as Signed English, Signed Japanese, Dutch in Signs, and the like. They use sign language vocabulary, but with the signs appearing in the order of the corresponding sentences from the spoken language. Sign-language-specific morphology is omitted, and morphemes that have no equivalent in the sign language are added, usually in the form of artificially invented signs for items such as articles, prepositions, inflectional endings, and so on. This may also involve the use of a manual alphabet which makes it possible to "fingerspell" words from the spoken language with letters represented by certain hand configurations. Depending on the degree of morphological difference between the signed and the spoken language, the utterance in the contrived hybrid system may be only a partial representation of the spoken language.

To the extent that deaf sign-language users are exposed to these systems, usually during their education, they become bilingual to some extent in the primary sign language and the contrived sign system. At this point, aspects of the contrived system may influence the primary sign language, so that, for example, a novel sign representing a particle or a preposition from the spoken language starts being used in the primary sign language as well. This has happened in the case of some of the question particles, as discussed in more detail in Chapter 140 on question particles. Contrived sign systems are used in many European countries, in Israel, in the English-speaking world including North America, Australasia, and parts of Africa (e.g. Uganda), as well as in some East Asian countries (Japan, Taiwan). Comparable systems are not in use to represent spoken languages in the South Asian subcontinent, in some other parts of Africa (e.g. Tanzania), in the Arabic-speaking world, and in Turkey.

5 Importance of sign language typology

Sign languages are of great theoretical interest to linguistics in general and to linguistic typology in particular. We are only just beginning to understand the similarities and differences across sign languages on the one hand and between signed and spoken languages on the other. The two chapters in this section present examples of the kind of linguistic diversity that can be found across sign languages. As more and more data from genealogically and geographically unrelated sign languages become available, we will gain a better understanding of the nature of typological variation across sign languages, and typological linguistics will be in a better position to make empirically substantiated generalizations across this visual-gestural linguistic type. Contrasting these results with spoken languages will then open up important new perspectives for our understanding of human language as a whole.

6 Transcription conventions

The examples from various sign languages presented in the two chapters in this section have been transcribed in a way that is widely used in sign-language linguistics. The transcription consists of glosses in capital letters representing the manual signs, above which nonmanual features are notated, such as facial expressions, head movements, and the like. The synchronization of nonmanual features with manual signs is indicated by the solid line separating the two, which is more than just a separator: its left-to-right positioning and length indicate when and for how long the manual and nonmanual gestures overlap.

The following abbreviations and symbols have been used in the examples:

INDEX-2	second-person pronoun
__neg	negative headshake
__q	interrogative facial expression
__pol-q	interrogative facial expression in polar questions
-	a hyphen is used when more than one word is needed to gloss a single sign, e.g. the sign NOT-WANT
+	a 'plus' symbol is used to gloss compounds, e.g. the compound sign HAVE+NOT-HAVE

Graphic representations of some signs have also been added for illustration, with arrows indicating movement of the hands.

139 Irregular Negatives in Sign Languages

ULRIKE ZESHAN

1 Defining the values

The map shows the distribution of irregular negatives in sign languages with respect to the number of such forms in each language. The following values are represented in the map:

○	1. No irregular negatives	1
○	2. One irregular negative	3
○	3. Some irregular negatives (2–5)	10
○	4. Many irregular negatives (more than 5)	21
	total	35

The term **irregular negative** refers to several types of forms that will be called **negative suppletion** and **negative derivation** here. They all have in common that they express negation in a nonstandard way (see Zeshan 2004b for details).

The canonical marking of negation in sign languages involves, on the one hand, **negative particles** that are independent signs and, on the other hand, **nonmanual negative signals** (head movements and facial expressions) that co-occur with the negated signs. Neither the negative particles nor the nonmanual negative signals affect the form of the sign or signs being negated. In example (1), from Indo-Pakistani Sign Language, there is a clause-final negative particle, and the nonmanual negative signal is a headshake (indicated by the line labelled "neg" on top of the sign glosses). The negated predicate 'interesting' takes the same form as it would in a positive clause.

(1) Indo-Pakistani Sign Language (own fieldwork)

neg

BOOK THIS INTERESTING NOT

'This book is not interesting.'

In the case of irregular negatives, on the other hand, the negation is part of the negated sign in one way or another, so that the negative form looks different from the positive form. These negative forms are called *irregular* because their number is always very small in any given sign language, with a maximum of around twenty-five items. Even if some of the negatives are consistently derived from positive forms in the same way, as happens in negative derivation, this is still a very limited process and quite different from the situation found in many spoken languages where a negative affix can be added to all verbs or predicates.

Negative clauses using irregular negatives usually cannot include any additional independent negative particle, but are otherwise no different from regular clause negation. Nonmanual negative signals combine with irregular negatives in the same way as with canonical negation (example (2) from South Korean Sign Language).

(2) South Korean Sign Language (K.-S. Byun, p. c.)

neg

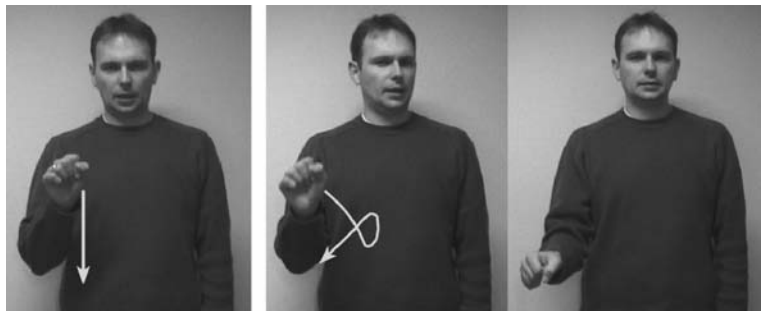
WAY NOT-KNOW

'(I) don't know the way.'

Negative suppletion involves negative forms that are entirely different from the corresponding positive forms. For example, the signs WANT and WANT-NOT in Russian Sign Language look entirely different and cannot be related to each other by any formational process (see Figure 1). **Negative derivation**, on the other hand, involves a set of negative forms that are derived from their positive counterparts in a regular way. There are two different types of negative derivational processes in sign languages. Negative derivation may take the form of internal modification of a sign, typically in the form of a modified movement pattern that is superimposed on a sign in a simultaneous way. For example, a set of eight negative signs is derived in DGS (German Sign Language) by adding a downward and diagonal inward-outward movement pattern to the positive sign (see Figure 2). The second type of negative derivation is sequential rather than simultaneous and is more comparable to negative affixing in spoken languages. For example, in a number of sign languages an outward twist of the wrist together with an opening of the hand is added to a positive sign to form a negative. The negative sign always ends in a position with the palm

WANT WANT–NOT

Figure 1 (Russian Sign Language)



CAN CANNOT

Figure 2 (German Sign Language)

LIKE LIKE–NOT

Figure 3 (Ugandan Sign Language)

KNOW–BAD

Figure 4 (Hong Kong Sign Language)

facing forward and all fingers extended. Figure 3 shows an example from Ugandan Sign Language. Another type of irregular negative, involving a negative handshape, is much less common across sign languages than negative movement patterns. A handshape with only the little finger extended, meaning BAD when used by itself, occurs in some East Asian sign languages and is used for negative derivation of both the simultaneous and the sequential type. The latter is exemplified in Figure 4, showing the sign KNOW-BAD 'not know' in Hong Kong Sign Language.

The typological distinction between a sign language with "some irregular negatives" and a sign language with "many irregular negatives" is not only a matter of the number of items, but correlates with another typological parameter. In the former type, all of the irregular negatives are suppletive forms, while the latter type involves both negative suppletion and negative derivation. This means that some of the negatives are formationally related by taking part in the same derivational process and thus forming a paradigm of derived negatives. For the purposes of this typology, it does not matter whether a language makes use of one or several derivational paradigms, how large the paradigm is, and which formational processes are applied to derive the negative forms. Finally, the category with "some irregular negatives" also includes a few borderline cases where no clear evidence has been found for the existence of a larger number of irregular negatives and/or the existence of a paradigm of forms related by a common negative derivational process.

2 Geographical distribution

Few generalizations emerge from the map in terms of areal patterns. Worldwide, irregular negatives are a common phenomenon across sign languages, with the majority of languages in the category having many irregular negatives. However, there is a striking absence of irregular negatives in the South Asian subcontinent, with a Pakistani variety entirely lacking such forms, and only a single irregular negative in Indian varieties of Indo-Pakistani Sign Language. Another sign language with just one irregular negative is Lengua de Señas Española (Spanish Sign Language). It is notable that the single irregular negative is a negative existential in both these cases. Moreover, the village-based sign languages are also strikingly poor in irregular negatives, with a single suppletive negative (aspectual NOT-YET) in Kata Kolok and only two irregular negatives in Adamorobe Sign Language.

For the most part, rather than being an areal phenomenon, the extent of irregular negativization seems at least partly to be linked to language families, although the notion of language family and genealogical grouping is itself unclear for sign languages (see introduction to the sign-language chapters). Both the British Sign Language (BSL) family and the French Sign Language (LSF) family have many irregular negatives. The BSL family also includes Auslan (Australian Sign Language) and New Zealand Sign Language. The LSF family also includes American Sign Language, Langue des Signes Québecoise, and Russian Sign Language. This group of sign languages also accounts for six of the eleven sign languages with a particularly high number of irregular negatives, amounting to ten items or more.

European sign languages as a whole predominantly have many irregular negatives with some negative derivation, and this is also true of their derivatives in other parts of the world, including North and South America, Australasia, Israel, and, indirectly via American Sign Language, Uganda. A particularly widespread type of negative derivation, involving an outward twist of the wrist and an open hand (see Figure 3), is also limited to European and European-based sign languages, occurring in Australia, New Zealand, France, Greece, Finland, Russia, Uganda, Brazil, Québec, the United States, and Israel. For Icelandic Sign Language, Irish Sign Language, and Lingua Italiana dei Segni, it has not been possible to establish the existence of a larger number of irregular negatives or a negative derivational process. The latter two sign languages do have individual forms that appear as part of larger paradigms in other sign languages, but in the absence of any other forms following the same patterns, they do not qualify as negative derivation in these sign languages.

There are a number of other such cases where signs that are part of a larger paradigm in other sign languages appear as individual irregular negative signs, and thus seem to have been borrowed. Borrowing and language contact do play a role in sign languages, but the contact situations that arise are much more complicated than is typically the case with spoken languages (see introduction to the sign-language chapters). It does seem that while individual irregular negative forms are readily borrowed almost on a worldwide level, derivational processes are not generally borrowed as a whole. This may account for a number of cases where we have a larger set of derivational irregular negatives in one sign language and one or two identical forms that are not part of a similarly large paradigm in another sign language. For example, Russian Sign Language has two signs IMPOSSIBLE and NOT-NECESSARY involving the same negative movement pattern as in German Sign Language (DGS), but lacks a more extensive paradigm (eight signs in DGS). American Sign Language has a larger paradigm of five signs with the same outward twist of the wrist, while Ugandan Sign Language has a paradigm with fewer members, including KNOW-NOT and LIKE-NOT, but not HAVE-NOT.

Sign languages in East Asia also have a large number of irregular negatives, but the most common derivational processes are different from those in European sign languages and their derivatives. There is extensive use of a negative handshape (see Figure 4) in Chinese Sign Language and, to a lesser extent, in Hong Kong Sign Language, which may be a related sign language. Reversal of movement for negativization is used in both these sign languages and in Nihon Shuwa (Japanese Sign Language). East Asian sign languages form language families of their own, but are not related to sign languages outside the region.

In addition to language families and language contact, there are many other factors that influence the number of irregular negatives. It seems that older sign languages are more likely to have many irregular negatives than sign languages that are very young or are the result of recent creolization of an indigenous and a foreign sign language. LSF, which is rich in irregular negatives, is one of the oldest documented sign languages. On the other hand, Thai Sign Language, which has only some suppletive irregular negatives, may have had a recent history of creolization due to influence from American Sign Language. Türk İşaret Dili (TID), however, seems to have had a longer history of its own, with evidence of advanced grammaticalization especially in the domain of negatives. This involves extensive negative cliticization, but only a few suppletive negative forms; many items that appear as irregular negatives in other sign languages participate in negative cliticization in TID instead, thus reducing the number of irregular negatives. Tanzania Sign Language, one of the young sign languages in the sample, does have several irregular negatives, partly due to the influence of gesture. If the local culture uses any negative gestures, these may be grammaticalized and incorporated into the sign language as signs, and this has happened with two negative gestures in Tanzania Sign Language. However, this sign language lacks any negative derivational processes.

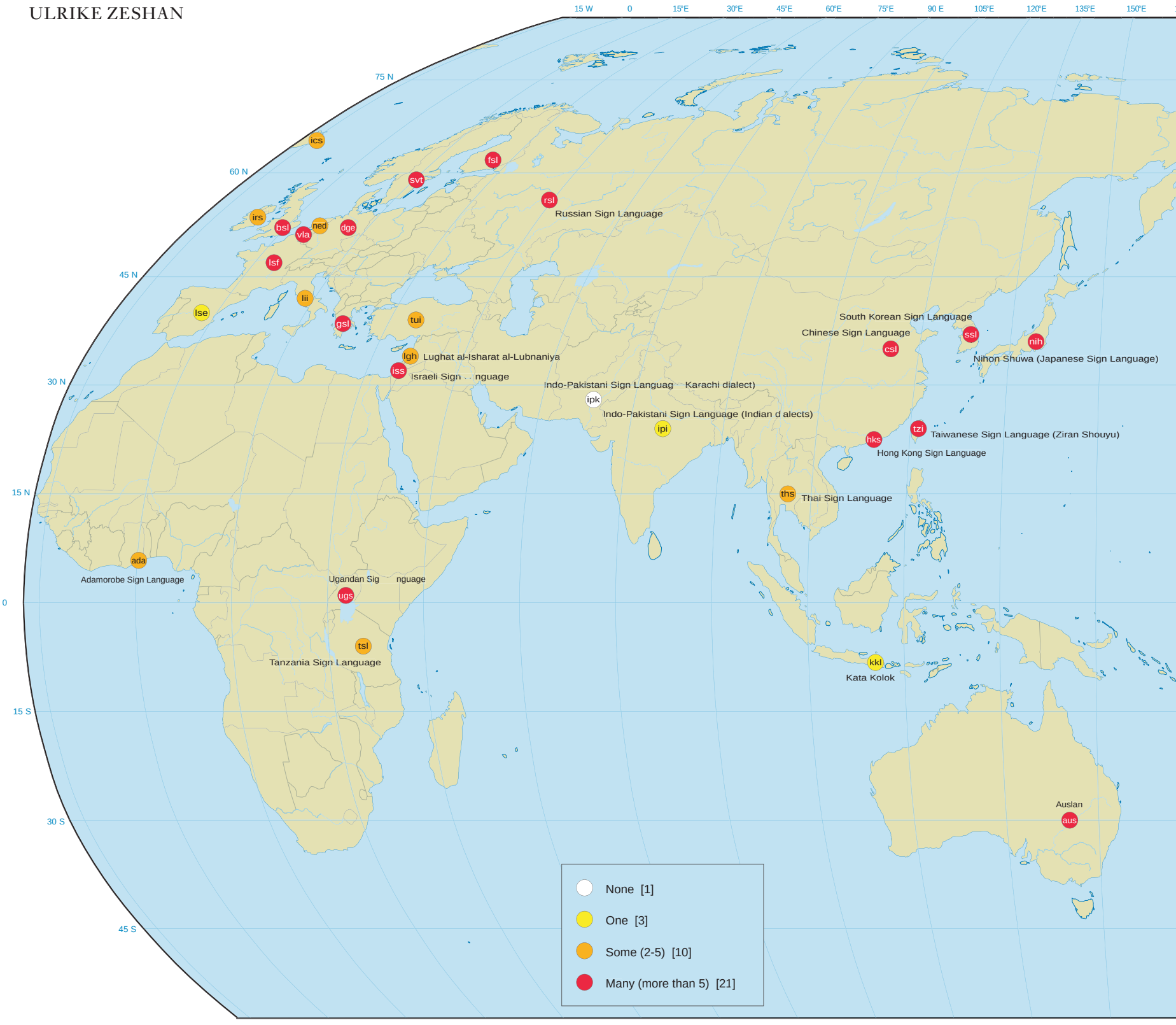
Irregular negatives typically occur in a number of specific semantic and grammatical domains. These include the following (with examples):

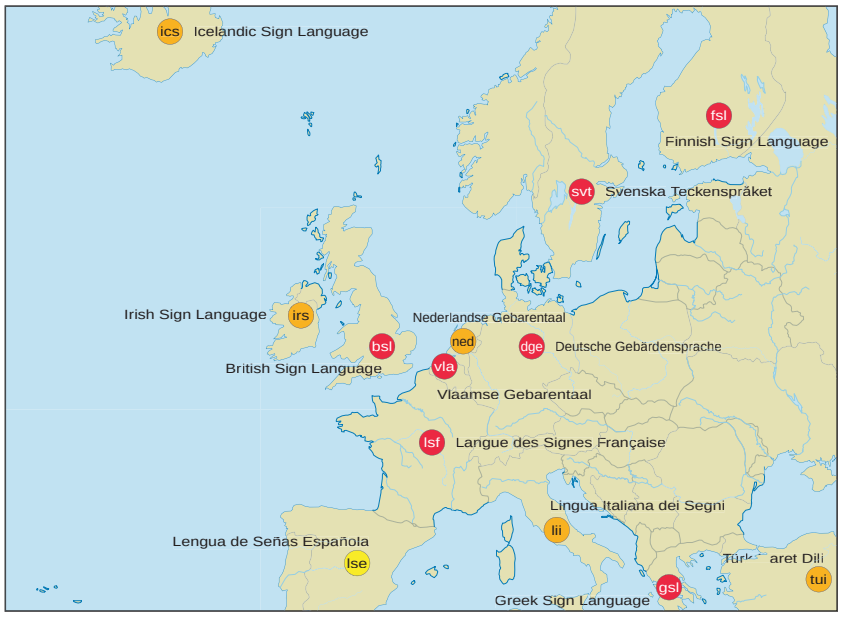
- cognition: not know, not understand
- emotional attitude: not want, not like, not care
- modals: cannot, need not, must not
- possession/existential: not have, not exist, not get
- tense/aspect: will not, did not, not finished
- evaluative judgement: not right, not possible, not enough

For sign languages that have only some irregular negatives (five or less), these are consistently drawn from a very restricted subset of this list, which are at the same time the most frequent irregular negatives across all sign languages. These include 'want', 'like', 'know', 'can', 'must', 'finish' (completive) and 'exist/have' (existential). The last and cross-linguistically most diverse category of "evaluative judgement" is the only domain that is not represented among the most frequent items.

139 Irregular Negatives in Sign Languages

ULRIKE ZESHAN





Europe

An example from Indo-Pakistani Sign Language

		neg
SPEAK	PUT:fwd	NOT-EXIST

"Voiceover has not been added."

Abbreviations:

:fwd	sign directed in space to forward location
neg	side-to-side headshake for negation

Note:
The sign NOT-EXIST occurs in Indian dialects of Indo-Pakistani Sign Language, but does not occur in some dialects from Pakistan.

65°E
180
165 W
150 W
135 W
120 W
105 W
90 W
75 W
60 W
45 W
30 W
75 N
60 N
45 N
30 N
15 N
0
15 S
30 S
45 S
Langue des Signes Québecoise
lsq
isl
International Sign
asl
American Sign Language
L ngua de Sinais Brasileira
lsb
Lengua de Señas Argentina
lsa
nzs
New Zealand Sign Language

140 Question Particles in Sign Languages

ULRIKE ZESHAN

1 Defining the values

For the purposes of this chapter, question particles are defined as signs whose main function is to indicate that an utterance is a question. Question particles are distinct from tag questions in that the latter involve an intervening intonational break, with only the tag marked by an interrogative facial expression (example (1)). They are also distinct from question introducers and pragmatic question markers. The latter mainly have a pragmatic function, carrying a more specific meaning such as expressing the speaker's attitude towards the content of the utterance, and they may also be prosodically detached from the rest of the utterance (example (2)).

(1) Auslan (T. Johnston, p. c.)

pol-q
CLASS CANCEL TODAY, RIGHT
'The class has been cancelled today, right?'

(2) Lingua Italiana dei Segni (Celo 1996: 143–4)

eyebrows up
body forward body back
ALWAYS (pause) QUESTION-MARK
'Will it be forever? (I don't know/I'm not sure/I don't believe it)'

To the extent that question particles may preferentially occur in particular subtypes of questions, such as confirmation questions, it can sometimes be difficult to distinguish them from pragmatic markers. Finally, question introducers are signs such as 'I ask you', 'do you mind if I ask', and the like, with a lexical rather than a grammatical meaning. Question particles, by contrast, have undergone a grammaticalization process and are thus more or less bleached of any original lexical meaning they may have had. It is not essential for the purposes of this chapter that a question particle should be obligatory in all questions, and indeed this is not the case in any of the sign languages in the sample.

The distribution of question particles is represented on the map by the following three categories:

○	1. No question particle	25
○	2. One question particle	9
○	3. More than one question particle	4
	total	38

The first category (no question particles) also includes sign languages for which no clear evidence for question particles has been found so far. In particular, in some sign languages there may be a process of incipient grammaticalization of question particles, but the evidence is not conclusive to speak of bona fide question particles. Similarly, the second category also includes sign languages for which no clear evidence of more than one question particle has been found so far.

Question particles in most sign languages are used in polar questions (or "yes/no-questions") only, with very few sign languages (e.g. Finnish Sign Language) allowing question particles to occur in content questions (or "wh-questions") as well. No question particle is used in content questions only, and no sign language has a distinction between a separate polar question particle and a separate content question particle. However, several sign languages do have more than one polar question particle. For example, Hong Kong Sign Language distinguishes between an existential question particle (also used with stative predicates) and a nonexistential question particle (used elsewhere, see example (6)). In Taiwanese Sign Language, one question particle has arisen because of influence from the surrounding spoken language while another one (see example (5)) is a form native to the sign language.

Examples (3) to (6), and the associated Figures (1) to (3), show examples of question particles in several unrelated sign languages. Question particles are almost always clause-final; in a few cases they occur either clause-initially or both initially and finally. They may be monomorphemic, like PALM-UP in Finnish Sign Language (example (4)), or morphologically complex, like the existential question particle in Hong Kong Sign Language, a compound of the signs HAVE and NOT-HAVE (see Figure 4).

(3) Lengua de Señas Española (Grupo de Investigación sobre Lengua de Señas, p. c.)

pol-q
nod
INDEX-2 IN SCHOOL DEAF YES-NO
'Do you go to a deaf school?'

(4) Finnish Sign Language (L. Savolainen, p. c.)

lowered brows
head tilt
PAPER WHERE PALM-UP
'Where can I find some paper? /Where is the paper?'

(5) Taiwanese Sign Language (W. Smith, p. c.)

pol-q
EAT FINISH HAVE-NOT-HAVE
'Have (you) eaten?'

(6) Hong Kong Sign Language (G. Tang, p. c.)

q
NOW TAKE-PHOTO GOOD-BAD
'Shall we take photos now?'

YES–NO

Figure 1 (Lengua de Señas Española)

PALM–UP

Figure 2 (Finnish Sign Language)

GOOD–BAD

Figure 3 (Hong Kong Sign Language)

HAVE+NOT–HAVE

Figure 4

2 Geographical distribution

The most striking areal pattern is the very high incidence of question particles in East Asia. In fact, all sign languages in this region

have question particles. In addition, the four sign languages where the existence of more than one question particle could clearly be established are also in this region (Hong Kong, mainland China, Taiwan, South Korea). All sign languages in the region use question particles in polar questions only. The sign languages used in South-East Asia and in the South Asian subcontinent do not have any question particles.

One cannot fail to notice that the high incidence of question particles in East Asia corresponds to a prevalence of question particles in the spoken languages in the region, such as Japanese *ka* or Mandarin *ma* (see Chapter 116). Although the sign language particles have quite a different status by virtue of being nonobligatory or even rare in the language, this distribution raises the question of a possible structural influence from the spoken languages on the sign languages. In fact, two question particles are known to have arisen through spoken-language influence via a signed representation of the spoken language (see introduction to the sign-language chapters about such sign systems). Thus the question particle in Nihon Shuwa and one of the question particles in Taiwanese Sign Language first arose via Signed Japanese and Signed Mandarin. On a continuum of registers between Signed Japanese and Nihon Shuwa on the one hand, and Signed Mandarin and Taiwanese Sign Language on the other hand, these question particles are still mainly used in registers closer to Signed Japanese/Signed Mandarin, but seem to be spreading into the native sign languages as well.

Taiwanese Sign Language has another question particle which is native to the sign language and is a grammaticalized contraction of the sign HAVE and its suppletive negative NOT-HAVE (see Chapter 139 for negative suppletion). In fact, this has effectively become a single sign in the present-day language. Moreover, there is a less grammaticalized layer of constructions used for forming polar questions that employ the same structure of positive and suppletive negative forms, such as CAN+CANNOT, KNOW+NOT-KNOW and WANT+NOT-WANT. However, these are not fused to the same extent as HAVE-NOT-HAVE. Similarly, Hong Kong Sign Language also has a question particle HAVE+NOT-HAVE (see Figure 4), but with a lesser degree of fusion and the two signs clearly distinct at present. All these constructions are strikingly similar to the Sinitic so-called "A-not-A construction", which is used in questions in the same way. In fact, we even find constructions that are exactly parallel to what we find in the sign languages, including the use of suppletive negatives. For example, spoken Cantonese uses an A-not-A construction consisting of a positive 'have' and a suppletive negative 'not-have'. Since we do not find such constructions and question particles in sign languages in other parts of the world, there is a strong possibility of a shared areal feature in this region.

Another strong argument in favour of spoken-language influence on the use of question particles in this region comes from three related sign languages in Japan, Taiwan, and South Korea. Despite this relationship, the question particles in these three sign languages are different from each other and show a different kind of spoken-language influence in each case. The signed equivalent of the Japanese interrogative *ka* has been entering Nihon Shuwa via Signed Japanese, while in Taiwanese Sign Language the use of forms following the Sinitic A-not-A construction patterns with its mainland Chinese counterparts rather than with its Japanese relative. Finally, the use of question particles in South Korean Sign Language correlates with three distinct formality levels. At the least formal level, no question particle is used, but a slightly more formal level includes a question particle. In very formal situations, a different, honorific particle is sometimes used which corresponds to the Korean politeness particle *yo*. Korean is known for its complex system of honorifics and speech levels, and the sign language partly follows this pattern with regard to question particles. No other sign language in the sample has a distinction in question particles that makes reference to an honorific category.

There is of course no necessary link between the occurrence of question particles in signed and spoken languages in the same region, as is evidenced by examples such as Thailand (question particle in spoken Thai, but no clear evidence for a question particle in Thai Sign Language) and the United States (question particle in American Sign Language, but not in spoken English).

European sign languages and their derivatives in the Americas and Australasia are generally poor in question particles, with only a few exceptions in Finland, Spain, and the United States. Unless more instances of question particles are discovered in the future, it is hard to see any regional or genetic patterns here. Finnish Sign Language uses a sign with one or two open hands turned so that the palm faces upwards (see Figure 2), while Lengua de Señas Española (Spanish Sign Language) has an entirely different form (Figure 1). American Sign Language uses a sign that is derived from the shape of a question mark, a source that is also exploited in Türk İşaret Dili and Tanzania Sign Language. The PALM-UP sign corresponds to a gesture that is widely used in parts of Europe, especially in combination with a shoulder shrug, and whose core meaning seems to be an expression of uncertainty, translatable as 'I don't know', 'no idea', and the like. More or less grammaticalized forms of this gesture turn up as question words, question particles, and negators in various European sign languages, as well as in eastern Africa (Tanzania, Kenya) and in International Sign (a sign pidgin used in international settings; see the introduction to the sign-language chapters). It is often difficult to pin down the function of a given instance of PALM-UP, or even to establish whether it should be classified as a sign or a gesture. Sign languages where PALM-UP occurs, but is not clearly identifiable as a question particle, include Tanzania Sign Language, British Sign Language, and Nederlandse Gebarentaal.

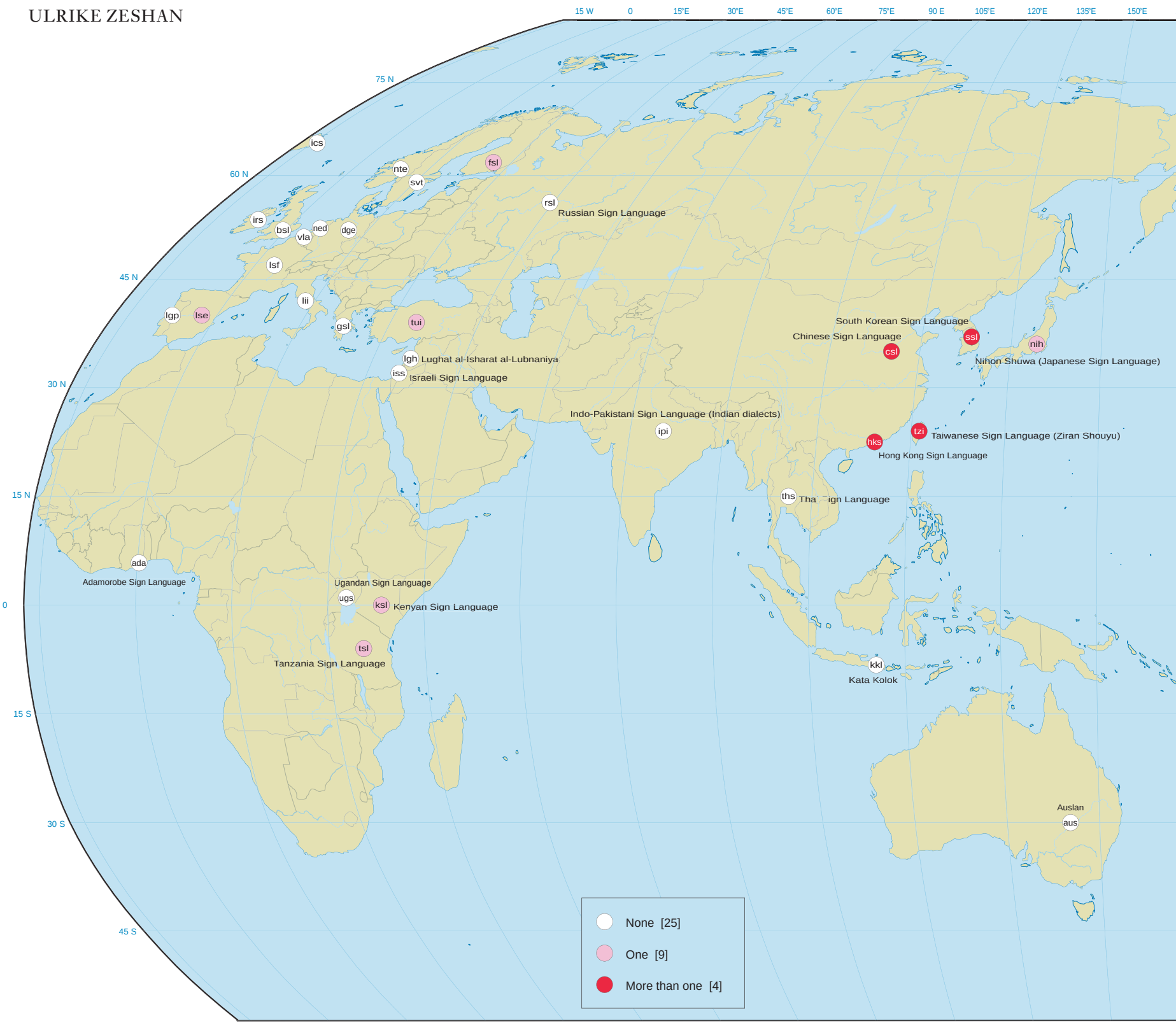
We do find more question particles in the Americas occurring outside the category of sign languages used in urban deaf communities. In Plains Indians Sign Language, a question particle is used in both polar and content questions and also has the semantics of a general interrogative in content questions, meaning 'what', 'where', 'when', 'why', and so on, according to the context. Urubú Sign Language uses a question particle in polar questions, unlike the spoken language of the tribe, which marks polar questions by intonation only. The two other village sign languages in the sample, from Bali and Ghana, do not have any question particles.

3 Theoretical issues

Typological studies on sign languages are of great theoretical interest to linguistic typology (see the introduction to the sign-language chapters). With respect to question particles, the relationship between signed and spoken languages and the grammaticalization of gesture are particularly interesting from a theoretical point of view (Zeshan 2004a).

140 Question Particles in Sign Languages

ULRIKE ZESHAN



Europe

An example from South Korean Sign Language

qu
EXIST/HAVE QU(ask)
"Do (you) have . . . ?"

Abbreviations:
QU (ask) question particle, derived from the sign ASK
qu non-manual marking in questions

Note:
The sign QU(ask) consists of two parts, shown in the second and the third picture. This question particle is used in semi-formal situations.

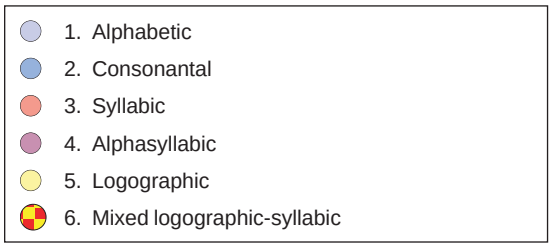
165°E
180
165 W
150 W
135 W
120 W
105 W
90 W
75 W
60 W
45 W
30 W
75 N
60 N
45 N
30 N
15 N
0
15 S
30 S
45 S
Langue des Signes Québecoise
lsq
Plains-Indians Sign Language
psl
asl
American Sign Language
Urubú Sign Language
usl
Língua de Sinais Brasileira
lsb
Lengua de Señas Argentina
lsa
nzs
New Zealand Sign Language

141 Writing Systems

BERNARD COMRIE

1 Defining the values

This map is concerned with the linguistic structure of different writing systems, in particular the unit of linguistic structure that is represented most directly in the writing system in question. Five basic systems can be represented, as given in 1–5 in the feature box; mixed systems are also possible, although only one, number six in the feature box, is included in the map.



In an **alphabetic** writing system, the basic unit of representation is the phoneme. Regular English spellings can be used to illustrate this type of writing system, e.g. *rat* /ræt/, *trim* /trɪm/, *spend* /spɛnd/. English can also be used to show that even in a writing system that is basically alphabetic, certain deviations from the strict principle of one-to-one correspondence of phoneme and grapheme may occur. Thus digraphs may represent a single phoneme, as with the initial two-letter sequence of *ship* /ʃɪp/, or more complex means may be used to represent certain phonemes, as when English uses the letter *e* after a consonant to indicate a different vowel quality of the vowel preceding that consonant, e.g. *fate* /feit/ versus *fat* /fæt/; these are regular phenomena in English spelling, even though they represent deviations from a strict alphabetic writing system. Primarily for historical reasons, English also has other, less systematic deviations from the alphabetic principle, as in *rough* /rʌf/, *bough* /bau/. Some other languages with an alphabetic writing system, such as Finnish, have an almost complete correspondence between grapheme and phoneme. Similar deviations from the strict system are also found with the other types discussed below.

A **consonantal** writing system is a variant of an alphabetic writing system with the added feature that only consonants, not vowels, are represented. For instance, the Hebrew consonant sequence *p–h* in (1a) is a well-formed orthographic word; note that Hebrew is written from right to left and that the Latin transliteration follows this order. Indeed, depending on which vowels it is pronounced with, it can be several different words, in Israeli Hebrew pronunciation *po* 'here' or *pe* 'mouth'.

(1) a. פה b. פֹּה c. פֶּה
hp hop hep

In many consonantal scripts, it is possible to add diacritics to indicate vowels, as in the forms (1b–c), respectively *po* and *pe*.

It will be convenient next to take **syllabic** writing systems, or syllabaries. In an ideal version of such a writing system, there is a distinct grapheme for each syllable of the language in question. The Japanese hiragana syllabary comes close to this ideal, as in the examples in (2). Note in particular that there is no part of the graphemes that can be identified as denoting the consonant phonemes /h/ or /n/, nor is there any part that can be identified as denoting the vowel phonemes /i/ or /e/; the level of representation is directly that of the syllable.

(2) ひ へ に ね
hi he ni ne

An **alphasyllabic** writing system, or alphasyllabary, falls in between the alphabetic and syllabic types. In attested alphasyllabic writing systems, the basic grapheme indicates a consonant, while a diacritic is added, or sometimes some other transformation applied, in order to indicate a combination of that consonant with a following vowel. The examples in (3) are from Thai, in which language the added diacritic can appear above, below, before, or after the consonant symbol, or some combination of these.

(3) ม มิ มุ มา เม เมอ
m mi mu maː meː məː

Note that, in contrast to the consonantal writing systems discussed above, in an alphasyllabic system the vowels must be indicated. (In practice, in alphasyllabic systems the consonant symbol without any diacritic usually has a so-called "inherent vowel"; in Thai, the leftmost example in (3) is often to be read /mo/.)

In all the writing systems described so far, the linguistic unit that forms the basic representational unit of the writing system has been identifiable in phonetic terms. In a **logographic** writing system, the basic unit of representation is the morpheme, i.e. a grammatical-lexical unit. Chinese examples, with the Mandarin pronunciation indicated using the pinyin transcription, are given in (4).

(4) 人 木 湖 壺
rén mù hú hú
'person' 'tree' 'lake' 'pot'

Since Chinese morphemes are nearly always a single syllable, one might wonder whether the Chinese writing system could not equally be described as syllabic. That this is not the case can be seen by comparing the written representations of the words for 'lake' and 'pot' in (4). They have identical phonetic representations, but are different morphemes with different meanings, and therefore have different representations; in a syllabary they would receive the same representation.

Most writing systems currently in use in the world fall into one or another of these five types. The main exception is Japanese, which is a **mixed** system, incorporating both logographic and syllabic principles. Roughly speaking, lexical morphemes (e.g. nouns, verb roots) are represented logographically, while grammatical morphemes (suffixes, particles) are represented by means of the hiragana syllabary.

(5) 人 ガ 犬 を 見た
hito ga inu o mi-ta.
man NOM dog ACC see-PST
'The man saw the dog.'

In (5), the nouns *hito* and *inu* and the verb root *mi-* are represented logographically, while the particles *ga* and *o* and the past tense suffix *-ta* are represented syllabically.

2 Geographical distribution

Representing the geographical distribution of the world's writing systems on a single map necessarily involves some oversimplification, in particular where more than one writing system is used in a given area, or where a particular writing system is used only in a very small area. For more detailed information, reference may be made to Daniels and Bright (1996). Moreover, the discussion in this chapter relates exclusively to the linguistic classification set out in §1, and disregards cultural links among writing systems. Thus, although most languages using Arabic script have a consonantal writing system, like that of Arabic itself, the versions of Arabic script used for writing Uyghur in northwestern China and for writing Kurdish represent all vowels, and are therefore considered alphabetic.

At present, the alphabetic writing system is clearly dominant in the world as a whole, being used in almost all of Europe (including

the Caucasus), northern Asia, Turkey and parts of Central Asia, insular South-East Asia and the Pacific islands, sub-Saharan Africa, and the Americas, in addition to pockets of use elsewhere, such as Korea and Vietnam. It is therefore in general easier to specify where the other writing systems are used.

The consonantal writing system is used throughout the Arabic-speaking world, and also for a number of other languages traditionally associated with Islam, in particular Persian, Pashto, and Urdu and other languages of Pakistan; it is also used for Hebrew, and occasionally for Tuareg and some other Berber languages. The consonantal writing system, in particular in its Arabic script version, was earlier more widespread; it is still used occasionally for such languages as Malay, Swahili, and Hausa, although it is now obsolete for Turkish and other Turkic languages.

The alphasyllabic writing system is found in three distinct parts of the world. The first is the "Indosphere", comprising nearly all of India, Nepal, Tibet, Bhutan, Bangladesh, Sri Lanka, Myanmar, Thailand, Laos, and Cambodia. Related scripts are also in occasional use in insular South-East Asia, from Sumatra through Java and Bali to South Sulawesi and parts of the Philippines. The second area is Ethiopia with an extension into Eritrea; in Ethiopia, the alphasyllabic Ethiopic script is used for most of the indigenous languages, including the national language Amharic. The third area is northern Canada, where an alphasyllabic script originally developed for the Algonquian languages Cree and Ojibwa is now also used for Athapaskan and Eskimo (Inuit) languages. The scripts of southern India and Ethiopia have numerous irregularities in the combination of consonant and vowel, thus moving them somewhat in the direction of a syllabary. The other alphasyllabic scripts of Asia and of northern Canada use almost exclusively regular combinations; those of northern Canada introduce the innovation of marking some vowels by rotating the consonant symbol, as in example (6) from Inuktitut (an Eskimo language).

(6)

ᐱ	ᐳ	ᐸ	ᑭ	ᑯ	ᑲ
pi	pu	pa	ki	ku	ka

Pure syllabic writing systems are rare in the modern world, even Japanese using its syllabaries only in conjunction with logograms. Two syllabaries developed in the first half of the nineteenth century are still in use, one for Cherokee, an Iroquoian language now spoken mainly in Oklahoma and North Carolina (USA), and one for Vai, a Mande language spoken in Liberia.

The Chinese script, used for all varieties of Chinese, is the only widespread purely logographic script in use today, although adaptations of Chinese script have been used for a number of neighbouring languages, and one is still current for Yi, spoken in southern China and belonging to the Tibeto-Burman family.

The mixed logographic–syllabic system used by Japanese has already been described in §1. Until recently, a similarly mixed logographic-alphabetic writing system was used by Korean. At present, only the alphabetic system is used in North Korea, and it is also dominant in South Korea.

3 The ancient world

The alphabetic writing system first arose with the Greeks, at a date that remains controversial but was probably in the eighth or ninth centuries BCE. Alphasyllabic writing systems are even younger. In the ancient world, then, we are dealing primarily with consonantal, syllabic, and logographic writing systems, including combinations thereof. The precise definition of the "ancient world" is of course open to interpretation, and in any event covers a period measured in millennia, but for practical reasons I have here taken the end point in the Old World to be the invention of the Greek alphabet; because of its completely separate history, Mesoamerican writing, though developing much later, is also included.

The oldest attested writing system is that of Sumerian, which arose in the late fourth millennium BCE; it is a mixed logographic-syllabic system. This Sumerian system was adopted and adapted directly or indirectly by a number of other languages of the region, most notably the Semitic language Akkadian. The Ancient Egyptian (hieroglyphic) script vies with Sumerian in terms of antiquity, and is a mixed logographic-consonantal script. In contrast to the general pattern among consonantal scripts, Egyptian script had a number of signs for sequences of two or three consonants (irrespective of intervening vowels), somewhat paralleling the use of the letter *x* to represent the two-phoneme consonant sequence /ks/ in Roman script.

The first purely consonantal writing systems arose in the Middle Bronze Age (first half of the second millennium BCE) among speakers of West Semitic languages, with an Egyptian model being the most probable main precursor. Among the earliest known purely syllabic scripts probably is Linear B, used for writing Mycenean Greek from the middle of the second millennium BCE.

Chinese script is logographic from its earliest clear attestations, around 1200 BCE, and has remained so ever since.

In Mesoamerica writing is attested from the second half of the first millennium BCE, with the Mayan script being by far the most widely attested and also the only one to have been substantially deciphered. The writing system is mixed logographic–syllabic.

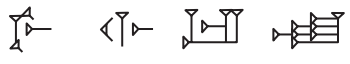
A number of features going beyond the above classification into the main types of writing systems characterize Sumerian, Egyptian, and Mayan as early writing systems. One is the rebus principle, whereby the symbol for a particular concept can also be used for homonyms; this principle was used especially in the creation of signs for abstract concepts and in the creation of phonetic (syllabic, consonantal) signs, as when the Sumerian sign for 'body', read *su*, is also used for the homonymous 'to replace'. The same principle is also found in the earliest Chinese writing, and is also used currently in transliterating foreign proper names into Chinese; it continues in modified form to play a role in that one character can be used as part of another character with the same sound, as in the examples in (7).

(7)

正	证
zhèng	zhèng
'straight'	'to prove'

Another principle is the use of so-called determinatives, namely signs that have no phonetic realization but serve to indicate the semantic field to which the word in question belongs. Finally, phonetic complements can be added to clarify part of the phonetic reading of a logographic sign or (in Egyptian) a two- or three-consonant sign. In the Akkadian example in (8a), a determinative 'woman' (the leftmost sign) is written before the proper name 'Shibtu' (written with the three syllabic signs *ši-ib-tu*). A parallel is again found in Chinese: the left-hand side of the second character in (7) is the character for 'speech', thus indicating the semantic field of the overall character. Akkadian example (8b) consists of a logogram (the first sign) meaning 'god' and pronounced *ilum*, followed by the syllabic sign for *lum*; since the same logogram can also be read *šamû* 'sky', the phonetic complement here helps to ensure the correct interpretation.

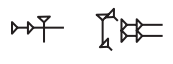
(8) a.



woman ši ib tu

'Shibtu (woman's name)'

b.

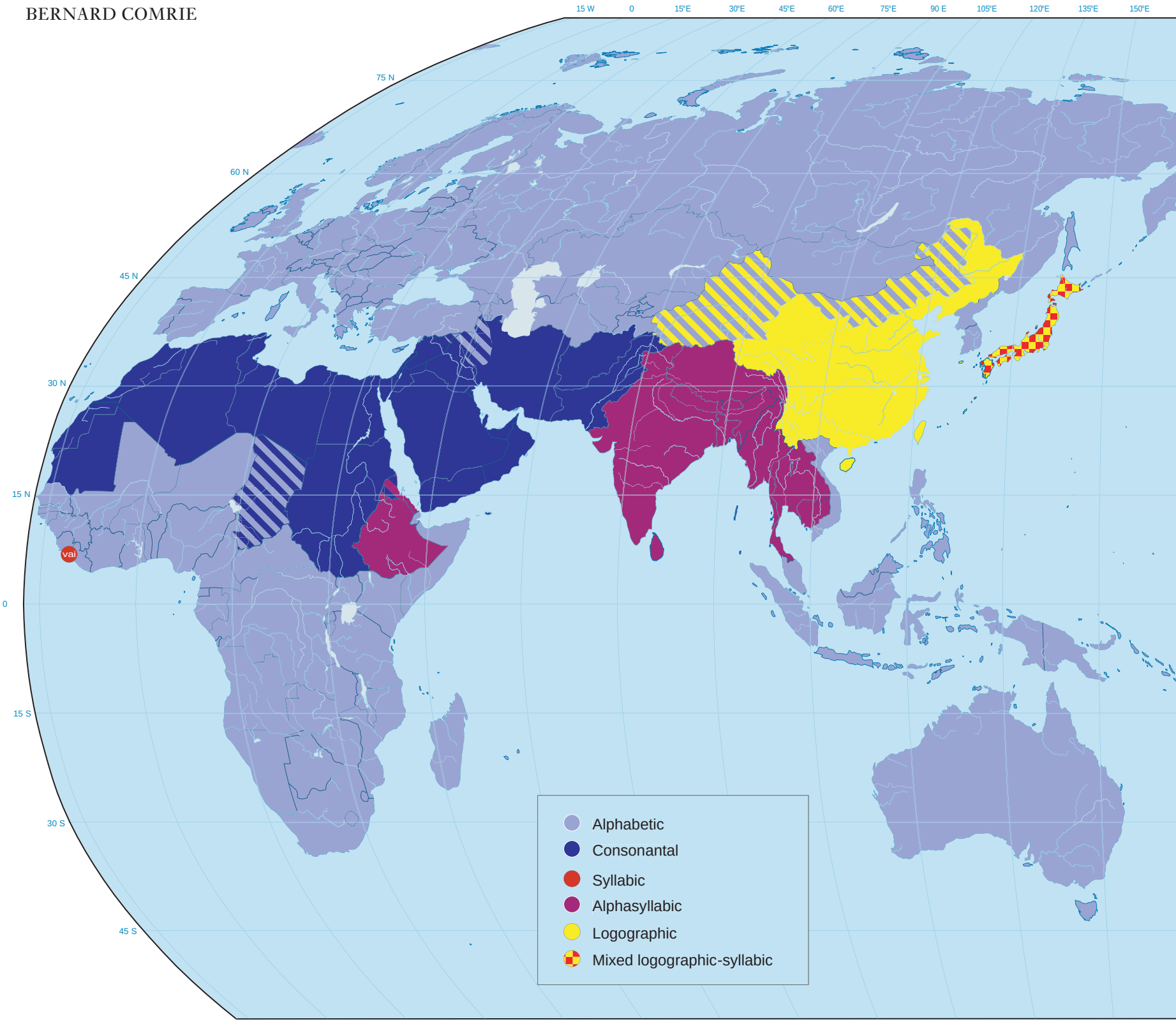


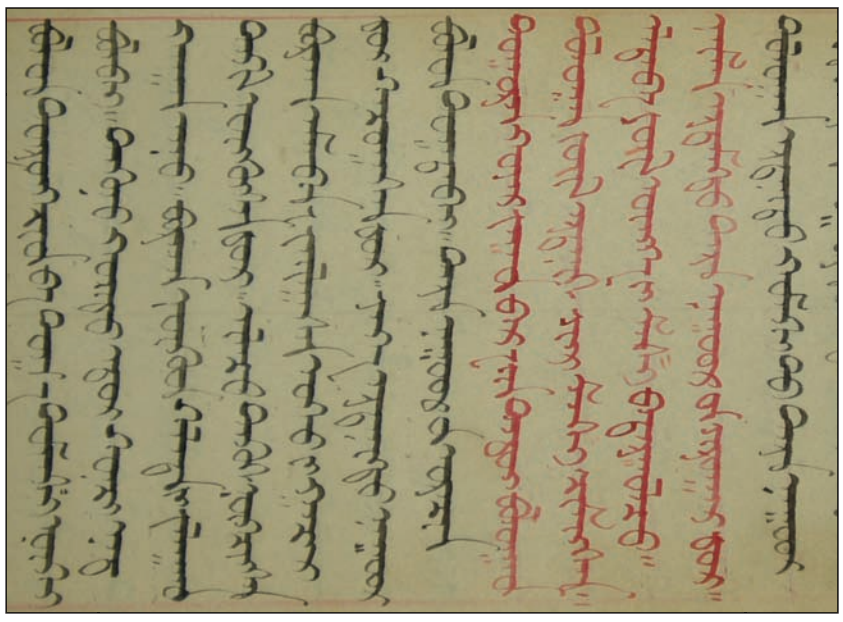
god lum

'god'

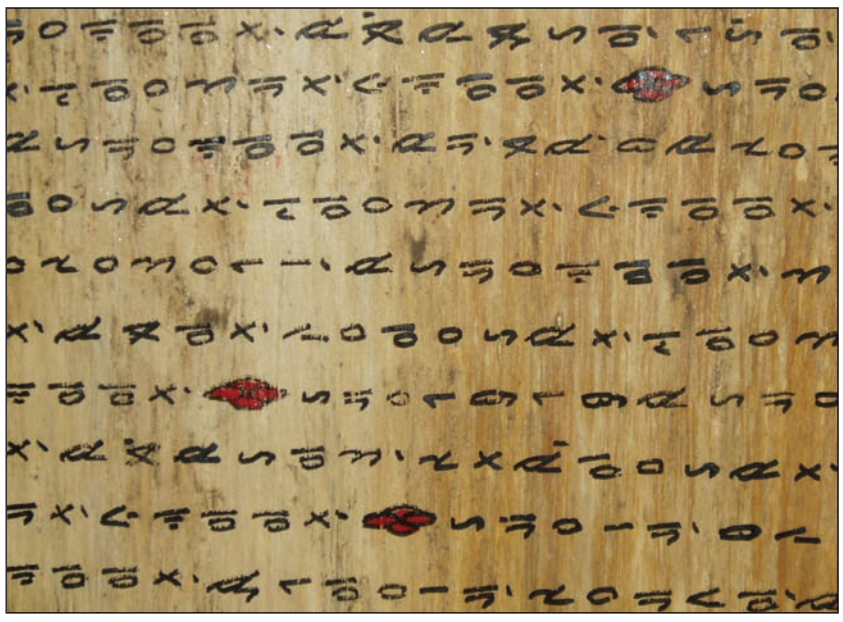
141 Writing Systems

BERNARD COMRIE

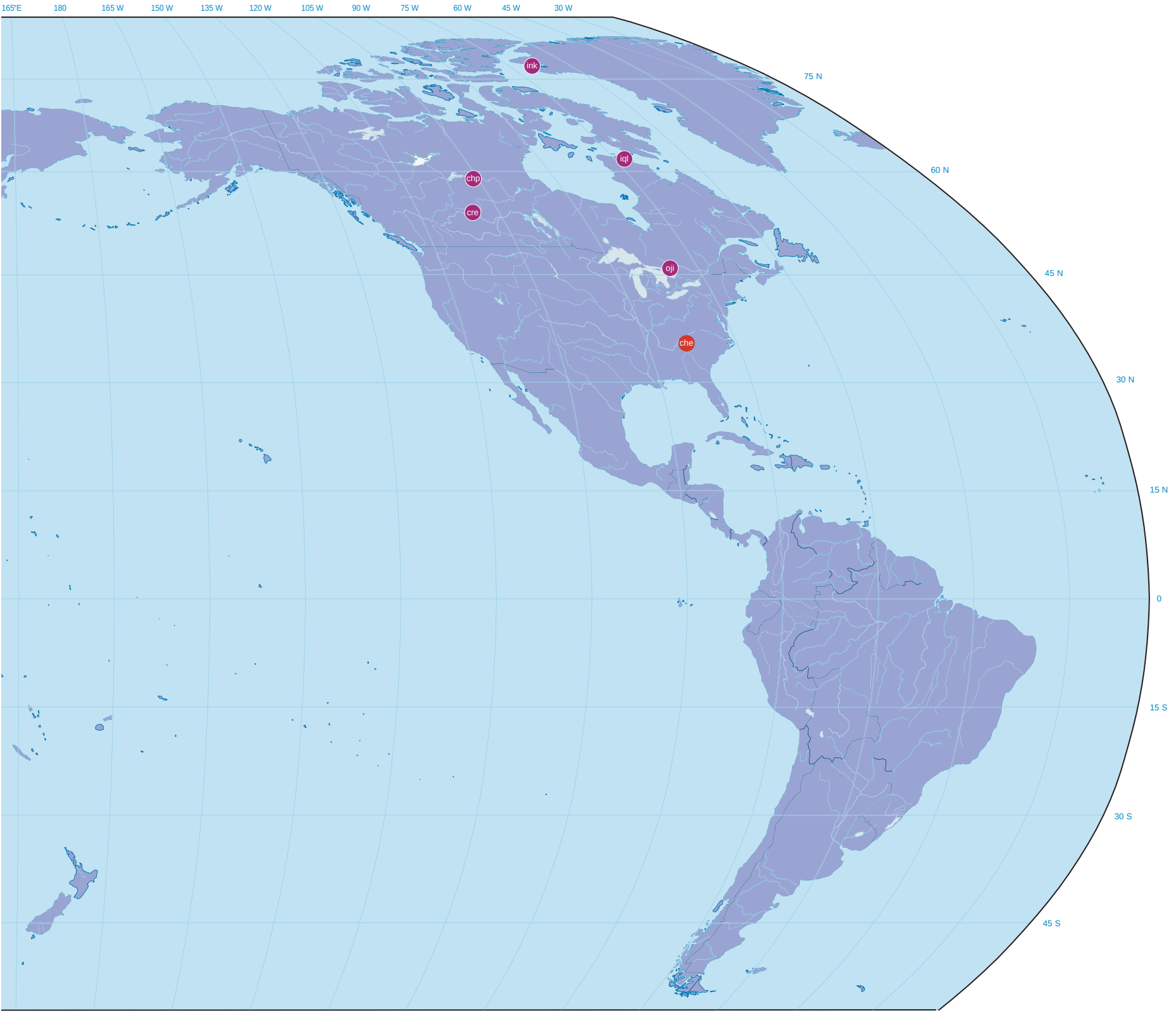


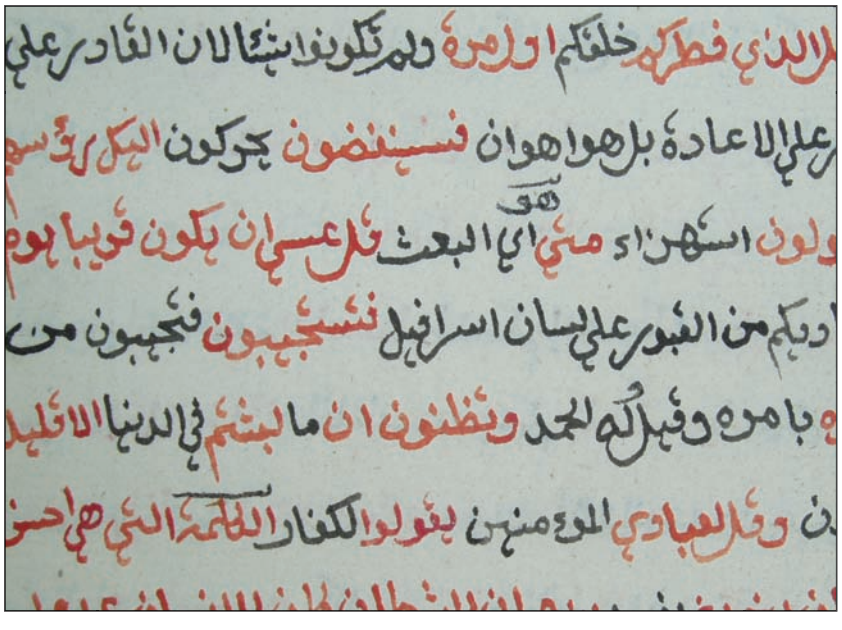


Alphabetic: Mongolian



Alphasyllabic: Batak





Consonantal: Arabic

女方怀孕后，先要请外公，不请外婆。他一个人带
来探望，可能是一种祈盼生男的旧俗。待孩子生下
外婆才来看女儿及外孙，来时带甜酒、肉、鸡蛋、鸡
如果男方家条件不好，就不给小孩办酒了，并没有
只有女人才能参加新生儿的庆典喜宴，亲家双
孩唱祝贺歌，赞美歌。小孩的背带另择时日自己家
会做，也可以请别人帮忙做。
小孩的命名需按照吴家的二十字派进行，如今也
再取小名。男人的发型发饰没有限制，长了就剪
长一些。女青年梳两辫子，或剪齐耳短发；生育以

Logographic: Chinese

142 Paralinguistic Usages of Clicks

DAVID GIL

1 Introduction

Clicks are acoustically salient speech sounds formed by the sudden opening of a closure resulting in a sharp inflow of air towards a secondary occlusion at the back of the mouth. As phonemes, the basic building blocks of words, clicks are very limited in their geographic distribution. Phonemic clicks occur only in southern Africa, plus a little pocket further to the north-east, in Tanzania; elsewhere in the world, they are absent. The limited distribution of phonemic clicks is displayed in Map 19.

However, click sounds are actually very common throughout the world, in a variety of usages which might be termed **paralinguistic**. A familiar example is the English sound which, when repeated, is spelled either as *tut tut* or as *tsk tsk*.

The paralinguistic nature of such clicks is reflected in a number of exceptional features. Phonetically, they involve sounds lying outside ordinary phonemic inventories. Grammatically, they are not integrated into morphological and syntactic structures. And semantically, they convey a very restricted range of meanings, some of which are associated with the expression of emotions.

The English *tut tut* is a dental click, often repeated two or more times, and is most commonly used to express feelings such as irritation, impatience, or disappointment. This usage may accordingly be characterized as **affective**; more specifically, as expressing negative affect. For some but not all speakers of English, the repeated dental click may also be used to express a very different range of emotions including amazement and appreciation; one context in which this occurs is that of men engaged in "girl-watching". This usage may be characterized as expressing positive affect. Whereas most speakers of English are consciously aware of the affective *tut tut*, there is an additional usage of a single dental click, typically immediately preceded by an opening of the lips, which occurs in generally subliminal fashion, without impinging on the consciousness of speakers and hearers: this is to mark the beginning point of a conversational unit, often in conjunction with the act of turn-taking. This usage can be readily observed worldwide on television news broadcasts such as CNN, in which the newscasters and reporters typically begin a stretch of speech with one of these clicks. In addition to dental clicks, some speakers of English make use of other clicks, either lateral or palatal, when addressing babies or domesticated animals, in order to attract their attention or to encourage them to engage in specific activities, such as, in the case of horses, running.

Paralinguistic clicks resemble other linguistic signs in that they are arbitrary and conventionalized. As such, they differ in both form and usage from one language to another. Accordingly, when speakers of one language travel to other countries, different usages of paralinguistic clicks may lead to misunderstandings. For example, an English speaker visiting Israel and asking a hotel receptionist if there are any rooms available might be answered with a single dental click. The English speaking guest may interpret such a click as expressing negative affect, and be bewildered by what seems to be an inappropriate and perhaps impolite response on the part of the hotel receptionist. However, in Modern Hebrew, unlike in English, a single dental click has little or no affective value; rather, it has a simple **logical** meaning, namely the expression of negation. Thus, the hotel receptionist was simply saying 'no'. In fact, the use of a dental click to express negation is characteristic not just of Hebrew, but of many Arabic dialects and other languages. However, in the San'ani dialect of Arabic, the dental click is used not for 'no', but rather for 'yes'; it thus expresses affirmation (Samia Naim, Martine Vanhove p. c.). Because such click sounds are paralinguistic, people sometimes leap to the assumption that they are universal, transcending ordinary linguistic boundaries. However, as examples such as these clearly show, paralinguistic click sounds vary from language to language and from place to place.

The paralinguistic usages of click sounds around the world were commented on by Charles Darwin (1872), in his treatise on *The Expression of the Emotions in Man and Animals*, where he refers to them as "clucks" and describes their occurrence in conjunction with particular gestures:

> "The throwing back of the head with a cluck of the tongue is said to be used as a negative by the modern Greeks and Turks [. . .] The Abyssinians, as I am informed by Captain Speedy, express a negative by jerking the head to the right shoulder, together with a slight cluck, the mouth being closed." (p. 274) "With the Hindoos Mr. H. Erskine concludes from inquiries made from experienced Europeans, and from native gentlemen, that the signs of affirmation and negation vary—a nod and a lateral shake being sometimes used as we do; but a negative is more commonly expressed by the head being thrown suddenly backwards and a little to one side, with a cluck of the tongue. What the meaning may be of this cluck of the tongue, which has been observed with various people, I cannot imagine." (p. 275) "According to three other observers, the Australians often evince astonishment by a clucking noise. Europeans also sometimes express gentle surprise by a little clicking noise of nearly the same kind." (p. 286)

2 Feature values

This map displays some of the ways in which the usages of paralinguistic click sounds vary across the world's languages. Specifically, the map portrays the distribution of two usages of paralinguistic clicks, affective, expressing either positive or negative affect, and logical, expressing either affirmation or negation. It does not show any of the other usages of clicks, such as for turn-taking or addressing babies and animals. In addition, it does not provide any information on the phonetic properties of the clicks, such as whether they are dental, lateral, or palatal.

1.	Clicks may express logical meanings ('yes' and/or 'no')	47
2.	Clicks may express affective but not logical meanings	71
3.	Clicks may express neither affective nor logical meanings	25
	total	143

In languages of the first type, there are paralinguistic clicks with logical usages expressing affirmation, negation, or both; in addition, such languages may or may not have clicks with affective usages. One example of such a language, mentioned above, is Hebrew. Another is Avar, in which a lateral click expresses negation (Konstantin Kazenin p. c.). In contrast, in Somali, a lateral click expresses affirmation (Giorgio Banti p. c.). In Mauritanian Fula, there is a contrast between a single dental click expressing affirmation, and two such clicks expressing negation (Alassane Dia p. c.); both of these clicks are produced with closed mouth and an acoustically salient coarticulation whose nature I was unable to identify. Some other languages with logical usages of paralinguistic clicks include Dyula (Cathy Crockford p. c.), Beja (Martine Vanhove p. c.), Sardinian (Michele Loporcaro p. c.), Georgian (Manana Bat-Hana p. c.) and Malayalam (M.T. Hany Babu p. c.).

In languages of the second type, there are paralinguistic clicks with affective usages but not with logical usages; in such languages, clicks may express positive affect, negative affect, or both. English is an example of this type. Another is Khumi, in which a single palatal click expresses negative affect; there is even a special verb *kədaŋ*, denoting the making of this sound (David Peterson p. c.). In Minangkabau, there is a contrast between a single dental click expressing negative affect, and multiple dental clicks expressing positive affect (own knowledge). Other languages of this type

include Swahili (Ellen Contini-Morava p. c.), Japanese (Daniel Long p. c.), Kalam (Andrew Pawley p. c.), O'odham (Virgil Lewis p. c.) and Huamelultec Oaxaca Chontal (Loretta O'Connor p. c.).

In languages of the third type, there are no paralinguistic clicks with either affective or logical usages. In such languages, there may be no paralinguistic clicks whatsoever, or there may be such clicks, but with other usages, such as for turn-taking or addressing babies and animals. Some examples of languages of this type include Ket (Ed Vajda p. c.), Lahu (Jim Matisoff p. c.), Mudburra (Patrick McConvell p. c.), Nez Perce (Noel Rude p. c.) and Hupda (Pattie Epps p. c.).

3 Geographical distribution

As is evident from the map, the distribution of paralinguistic clicks forms a geographical pattern on a global scale. In fact, the present map dovetails with Map 19 showing the distribution of phonemic clicks. Combining the two maps, languages may be characterized in accordance with the functional load of click sounds:

Map 19 types 2 and 6	phonemic clicks	
this map, type 1	paralinguistic clicks:	logical
this map, type 2	paralinguistic clicks:	affective
this map, type 3	paralinguistic clicks:	neither

In the above diagram, the arrow at the left represents decreasing functional load of clicks. The functional load is highest in languages with phonemic clicks, decreasing in the three types of languages portrayed in this map, in the order in which they are presented here.

Looking at the two maps, a striking pattern emerges: the functional load of click sounds decreases as one moves out of southern Africa, northwards through the rest of Africa and into Europe, and then eastwards across the Eurasian continent, south-east to New Guinea and Australia, and across the Bering Strait into the Americas.

As shown in Map 19, and mentioned once again at the beginning of this chapter, languages in which clicks are associated with the highest functional load, namely as phonemes, are found exclusively in southern Africa, with an additional enclave in Tanzania. A single, isolated quasi-exception to this generalization is provided by the Australian language Damin, in which phonemic click sounds occur in a special speech style associated with initiation rites.

Among the remaining languages, the ones in which paralinguistic clicks bear the greatest functional load are those of the first type, in which clicks are used to express the logical meanings of affirmation or negation. As is evident from the present map, these languages occur in a single large-scale contiguous region stretching across western and northern Africa, southern Europe, and southwestern and southern Asia: more or less from Senegal to Bengal. A single exception, which proves the rule, is provided by the creole language Sranan, which clearly reflects a recent trans-Atlantic migration from West Africa.

Within languages of this first type, finer distinctions not displayed in the map provide further support for the generalization that the functional load of click sounds decreases as one moves away from Africa. Languages with contrasting clicks for affirmation and negation, such as Mauritanian Fula mentioned above, are limited to Africa. In Eurasian languages of this type, there is at most one click expressing a logical meaning. In fact, within languages of this type, there is a large continuous band consisting of languages which, like Hebrew discussed above, make use of a dental click to express negation; this isogloss extends from Morocco in the west, across the Mediterranean and southern Europe, into the Middle East and parts of the Caucasus, and then further east, tapering out in the South Asian subcontinent.

Languages of the second type, in which paralinguistic clicks have affective usages but not logical ones, are widespread in many different parts of the world. From a geographical perspective, they would seem to constitute a kind of default background upon which other isoglosses may be superimposed in particular regions.

Languages in which paralinguistic clicks bear the least functional load are those of the third type, in which there are no paralinguistic clicks with either logical or affective usages. As is evident from the map, there are no such languages in Africa or for that matter in Europe; progressing eastwards, the first occurrences of such languages are in northern Siberia and in Myanmar. Such languages predominate in northeastern Asia and in North America; they also occur, interspersed with languages of the second type, in South-East Asia, Australia, and South America.

4 Theoretical issues

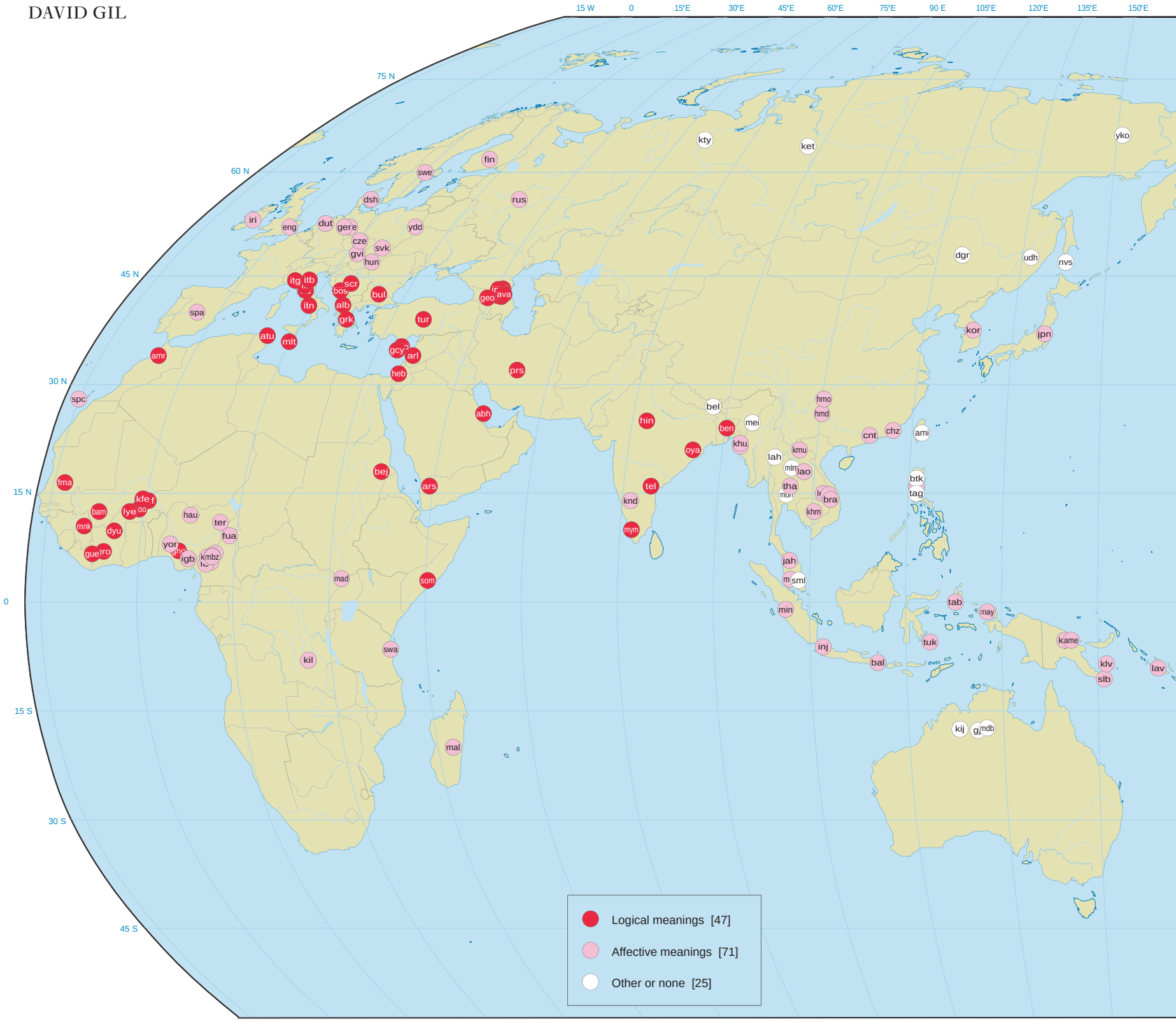
As this map and Map 19 show, the functional load of clicks decreases as one moves out of Africa, north, then east, and then south across the world. How can this remarkable pattern be explained? Coincidence is unlikely: the geographical patterns are too dramatic to be attributed to mere chance.

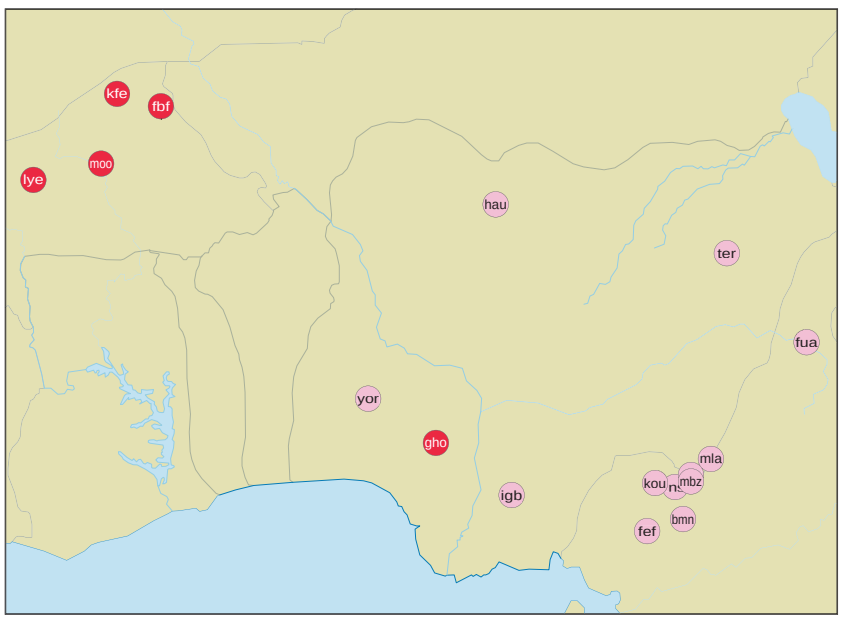
One cannot avoid the observation that the functional load of click sounds decreases more or less consistently as one follows the reconstructed trajectory of the ancient migrations which brought mankind out of Africa to populate the rest of the world. But what can this mean? Geneticists cite the genetic diversity of Africans to support the claim that mankind originated in Africa. By the same token, then, one could argue that the functional diversity of click sounds in Africa is evidence that click sounds originated in Africa and spread out from there to the rest of the world. However, even if one accepts both conclusions, it does not necessarily follow that the two events were contemporaneous, that is to say, that clicks originated with mankind, and accompanied man's earliest migrations out of Africa. Click sounds could have originated much later, or for that matter much earlier, as part of a pre-human protolanguage; at present we just do not know.

It is often assumed that large-scale geographical patterns such as those depicted in this map must reflect events in ancient prehistory: either inheritance from a common ancestral language, or else contact between ancient languages. This assumption presupposes that the feature in question is endowed with a sufficient degree of diachronic stability to enable geographical patterns, once they emerge, to be maintained over the course of time. However, there is at least some evidence suggesting that the various usages of click sounds may not be impervious to recent borrowing. For example, in southern Africa, recently arrived Bantu languages such as Xhosa borrowed phonemic clicks from the indigenous Khoisan languages. Similarly, the large-scale isoglosses of this map occasionally cross-cut closely related languages. For example, within Romance, Italian has logical usages of paralinguistic clicks while French does not, and within Slavic, Serbian-Croatian has logical usages of paralinguistic clicks while Czech does not. Likewise, within the Kuki-Chin-Naga branch of Tibeto-Burman, Khumi has affective usages of paralinguistic clicks while Meithei does not, while within the Aslian branch of Mon-Khmer, Jahai has affective usages of paralinguistic clicks while Semelai does not. Indeed, an epidemiologist might easily produce a map showing similar large-scale patterns reflecting the worldwide spread of, say, a virus within the course of mere years. What these observations suggest, then, is that at present we are still far from being able to determine whether the geographical patterns evident in this map are relatively old or comparatively recent, and hence even further from being able to come up with an adequate explanation for their existence.

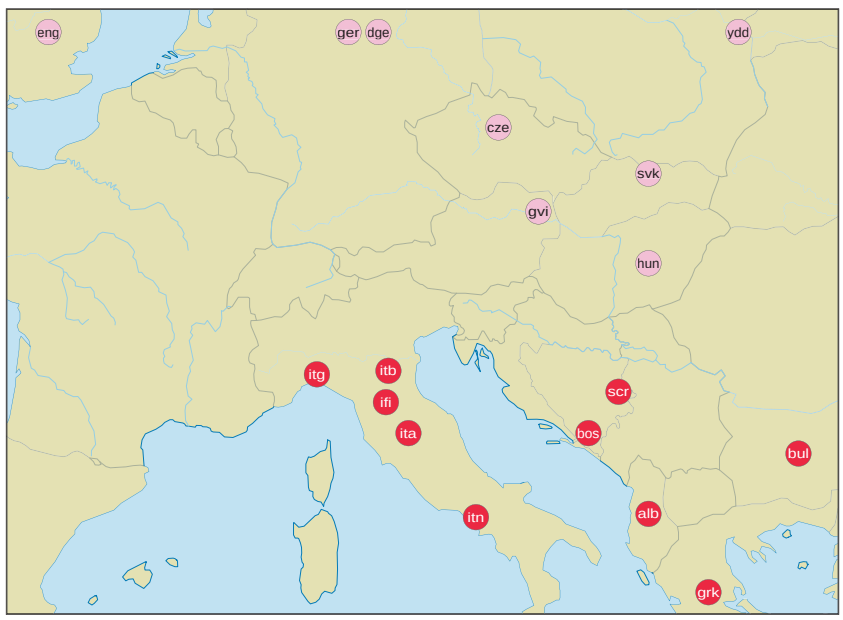
142 Paralinguistic Usages of Clicks

DAVID GIL

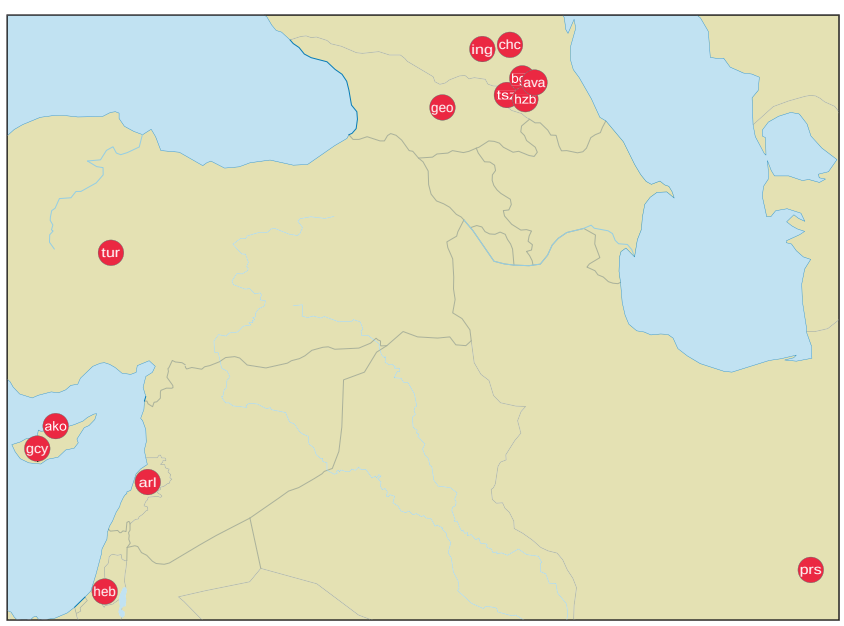




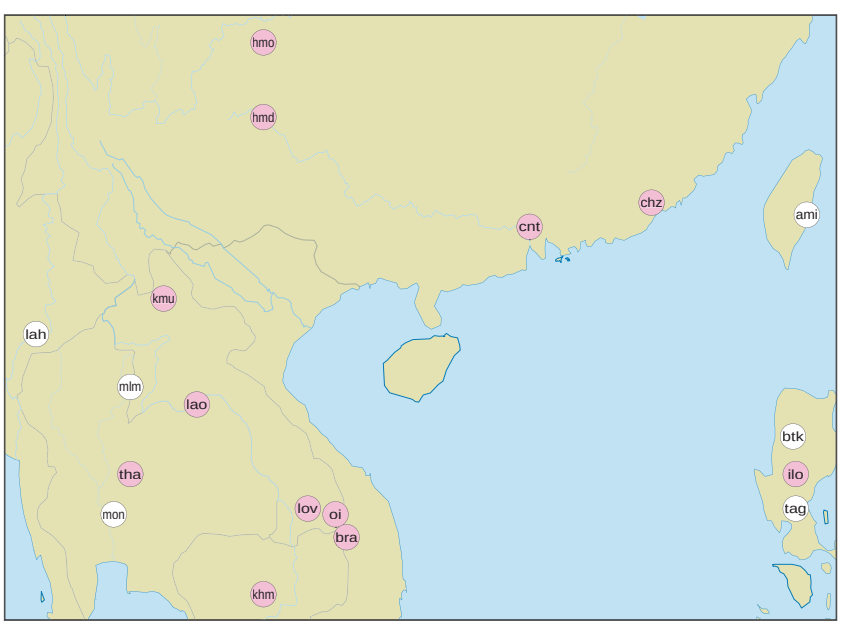
Nigeria and surrounding areas



central and southern Europe

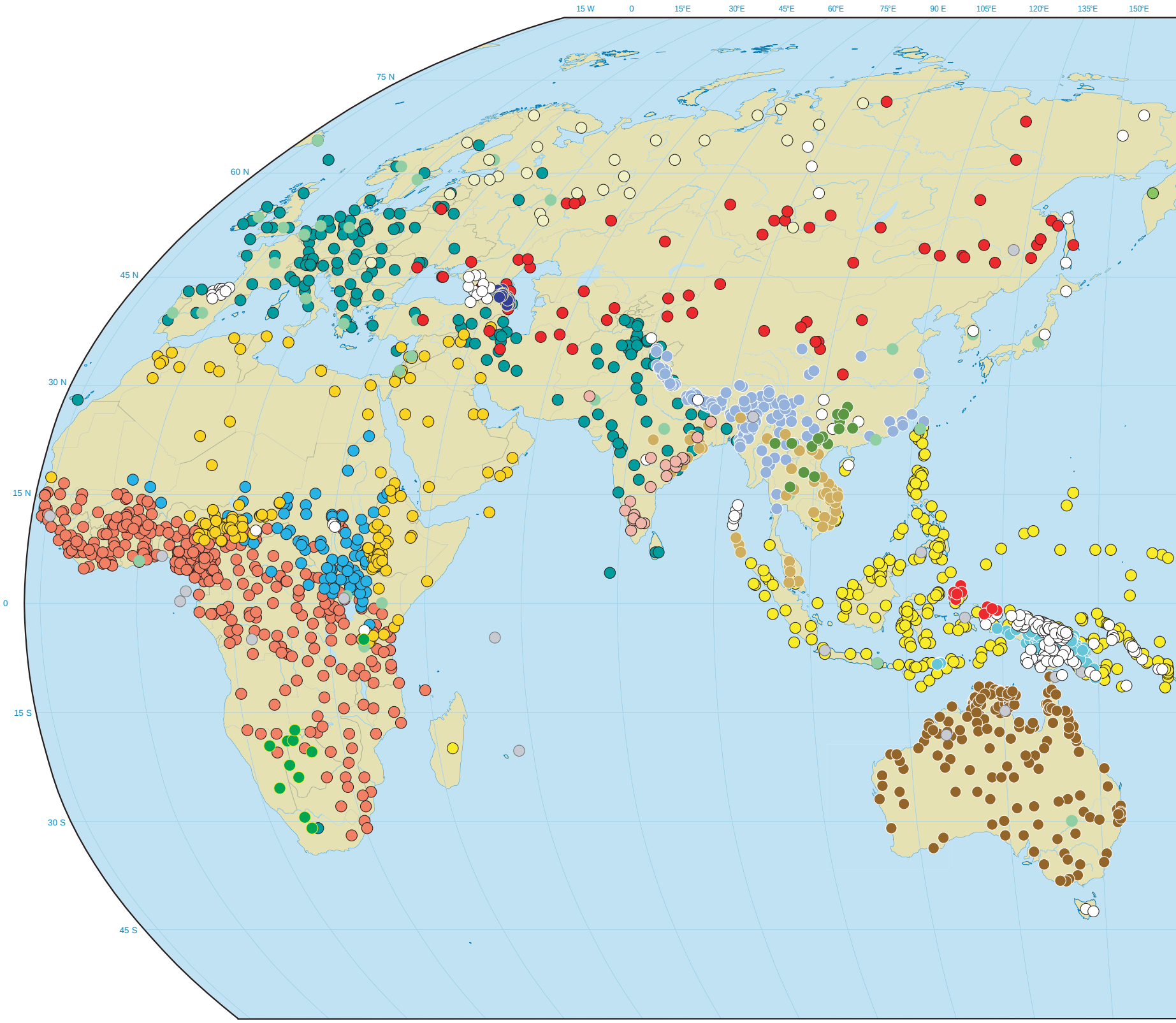


Caucasus and Middle East



northern South-East Asia

The World



Africa

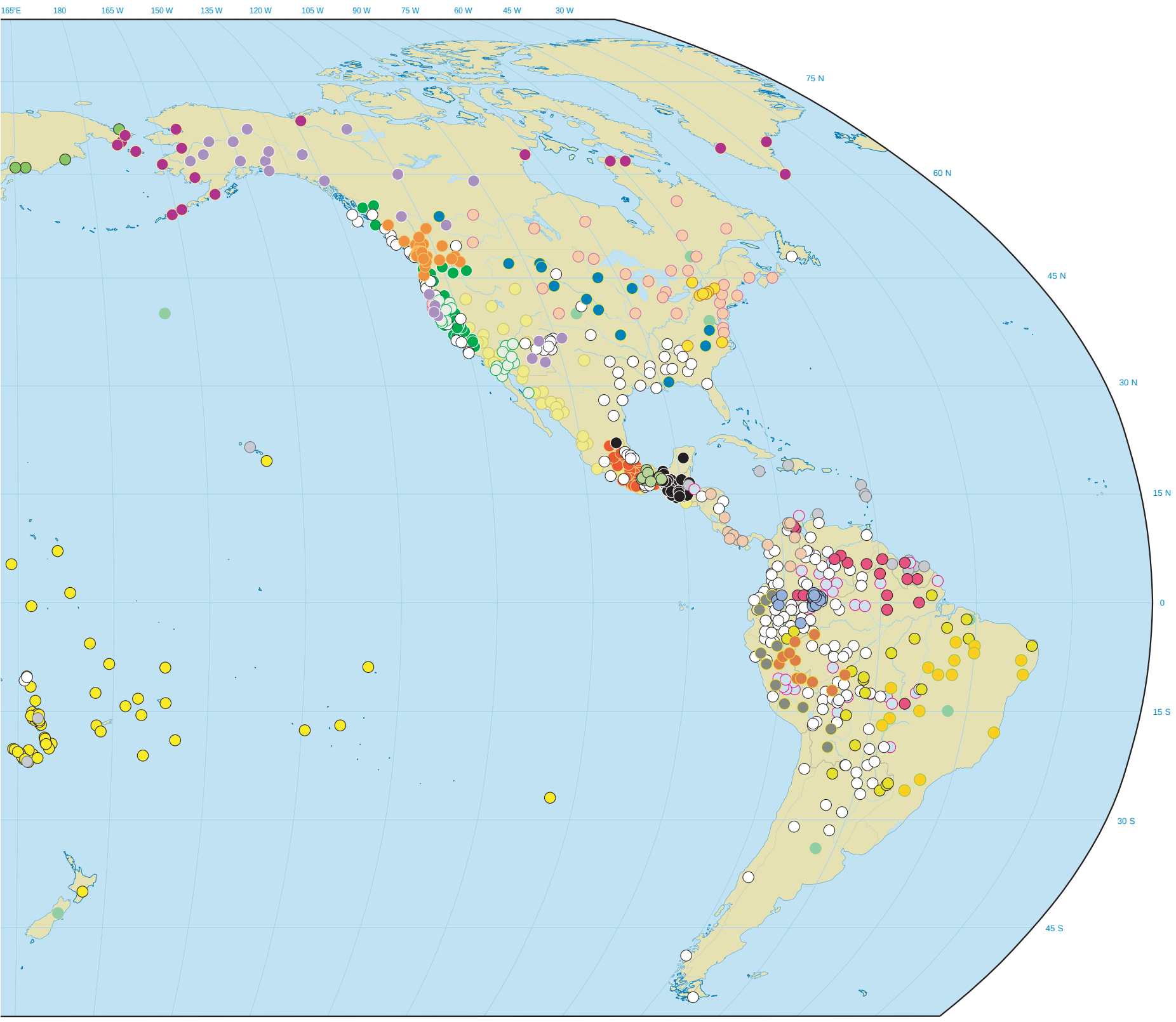
- Afro-Asiatic
- Khoisan
- Niger-Congo
- Nilo-Saharan

Eurasia

- Altaic
- Austro-Asiatic
- Chukotko-Kamchatkan
- Dravidian
- Indo-European
- Nakh-Daghestanian
- Sino-Tibetan
- Uralic
- Tai-Kadai

Pacific

- Australian
- Austronesian
- Trans-New Guinea
- West Papuan



North America

- Algic
- Eskimo-Aleut
- Hokan
- Iroquoian
- Mayan
- Mixe-Zoque
- Na-Dene
- Oto-Manguean
- Penutian
- Salishan
- Siouan
- Uto-Aztecan

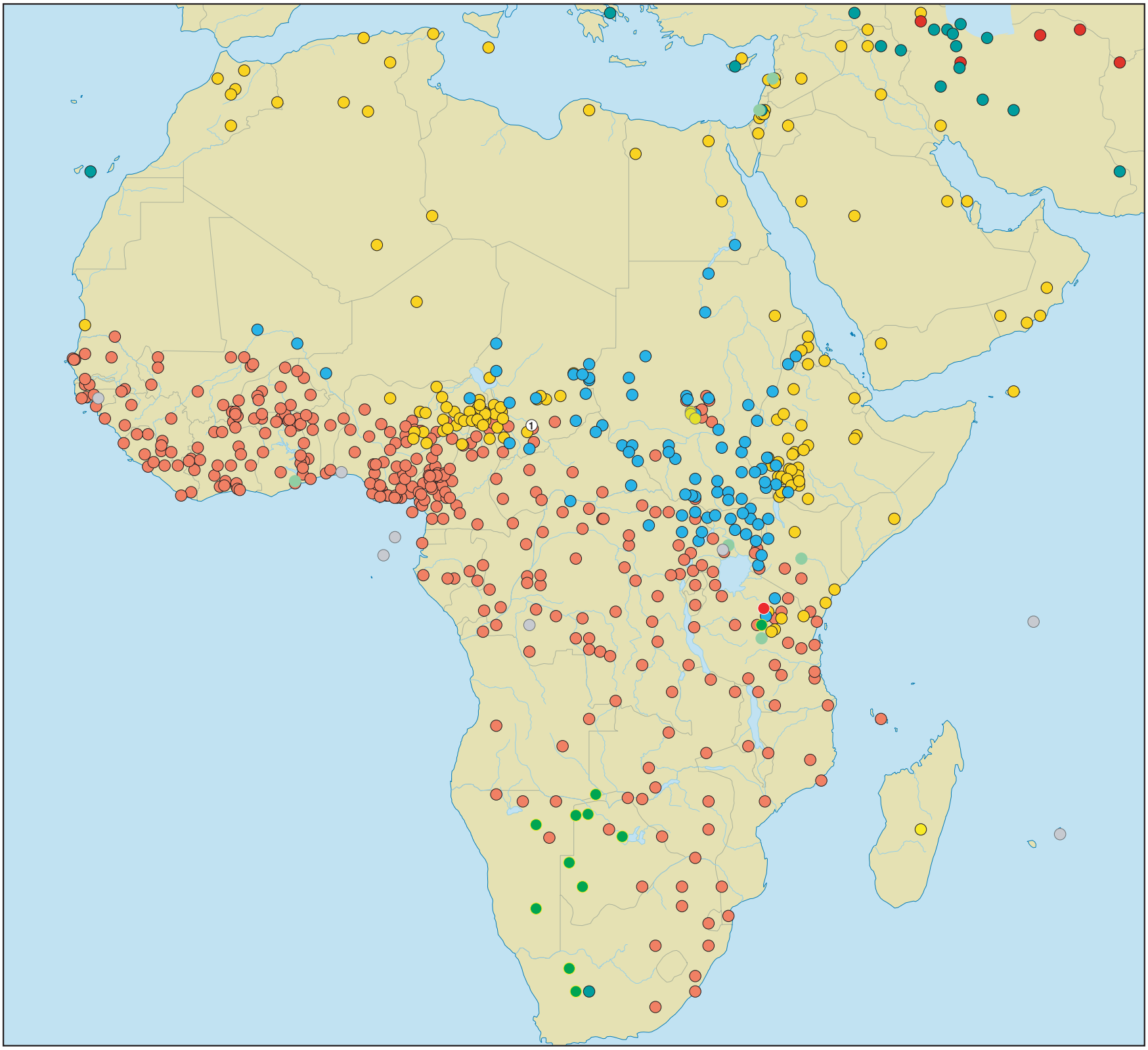
South America

- Arawakan
- Cariban
- Chibchan
- Macro-Ge
- Quechuan
- Panoan
- Tucanoan
- Tupian

Worldwide

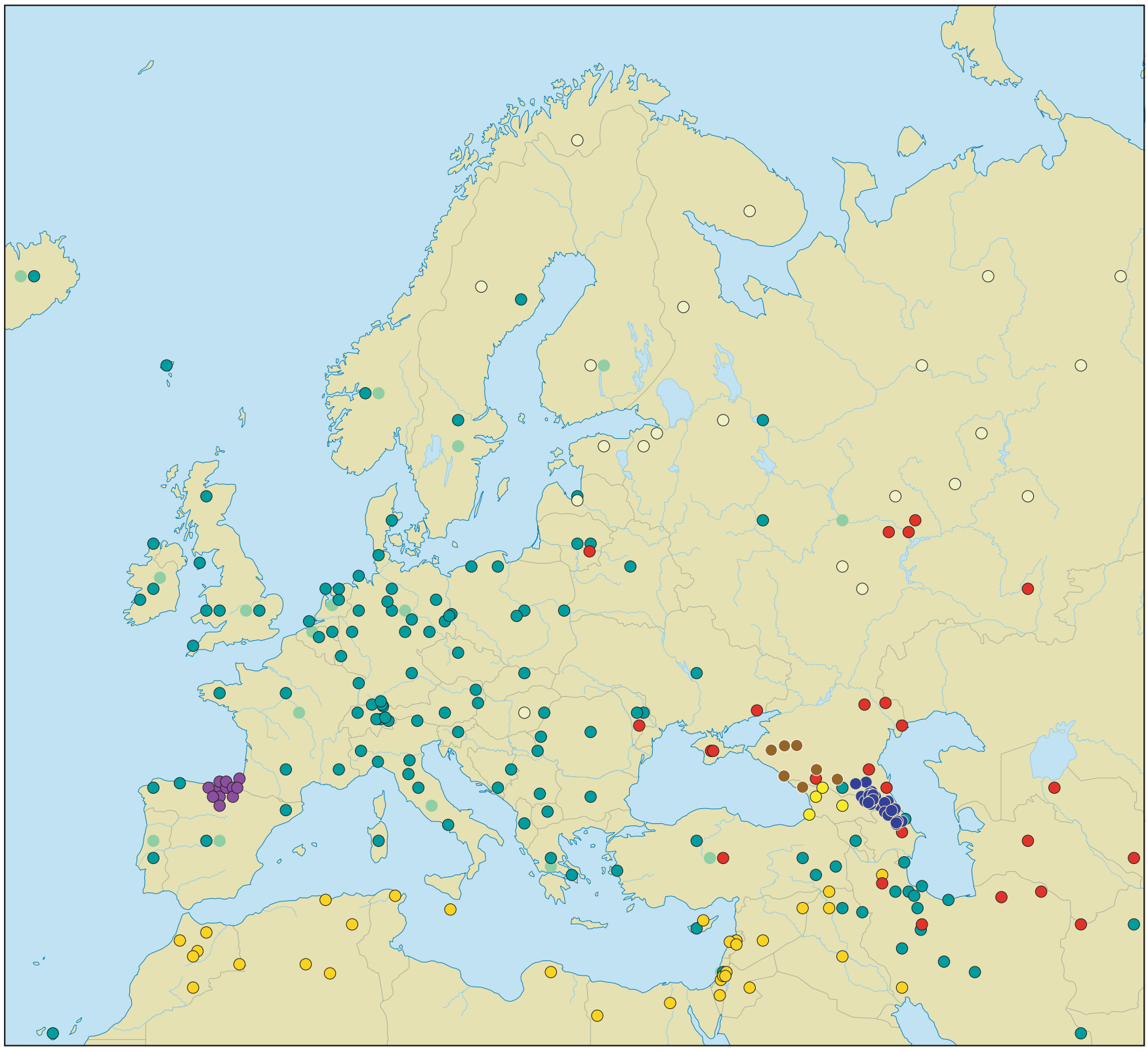
- Creoles and Pidgins
- Sign Languages
- Other Families and Isolates

Africa



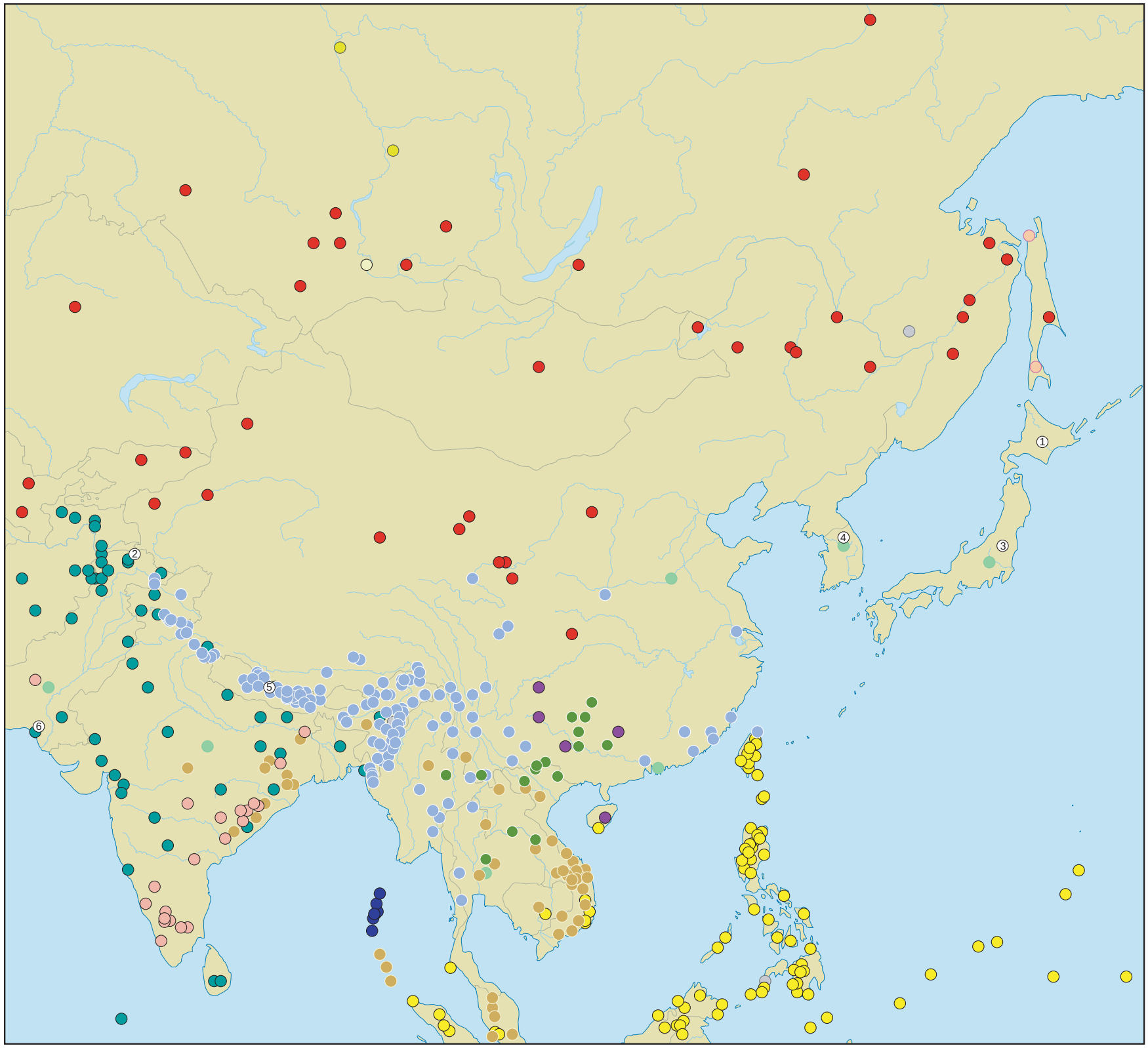
- Afro-Asiatic
- Altaic
- Austronesian
- Creoles and Pidgins
- Hadza
- Indo-European
- Kadugli
- Khoisan
- ① Laal
- Niger-Congo
- Nilo-Saharan
- Sign Languages

Europe



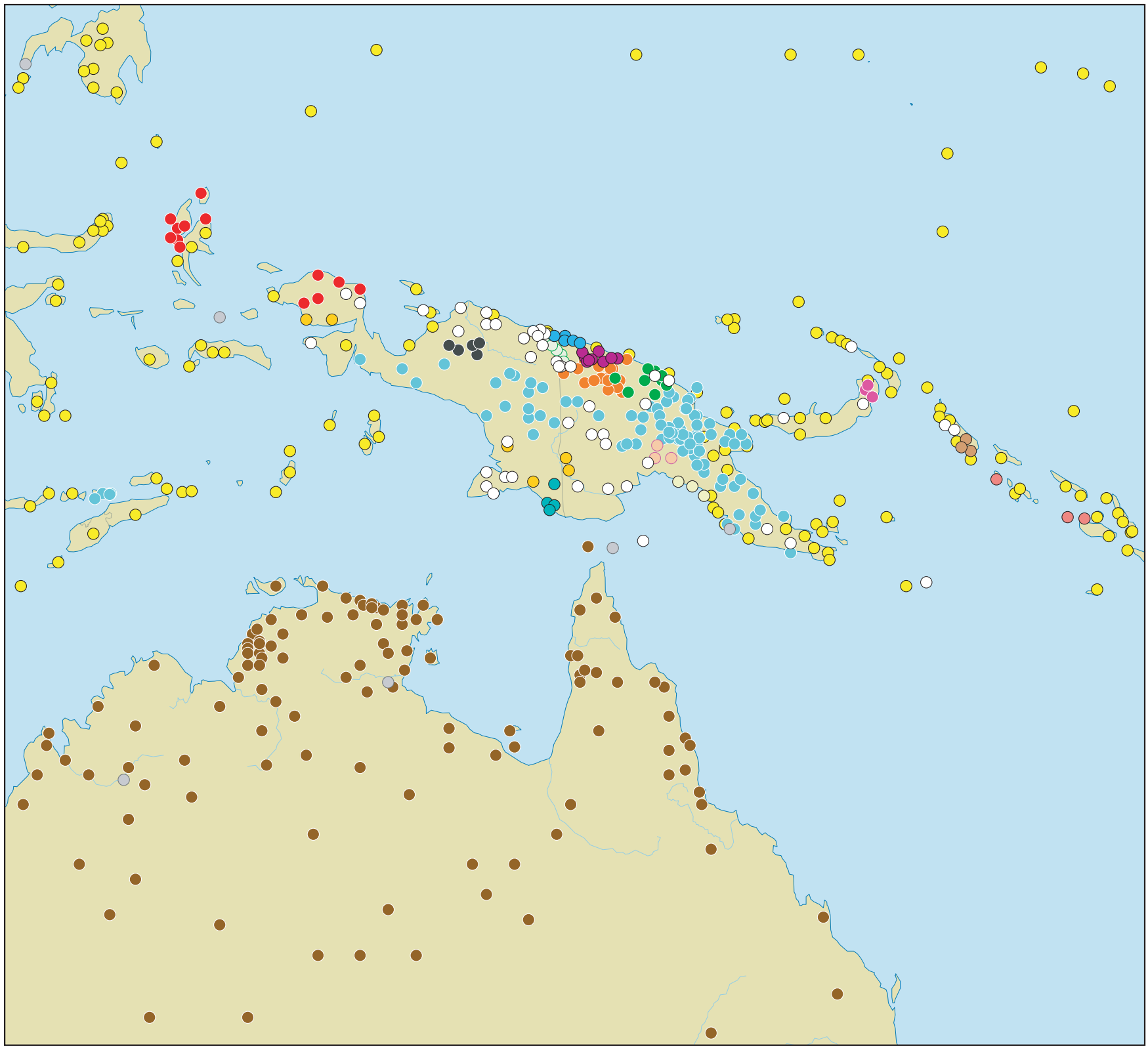
- Afro-Asiatic
- Altaic
- Basque
- Indo-European
- Kartvelian
- Nakh-Daghestanian
- North-West Caucasian
- Sign Languages
- Uralic

Asia



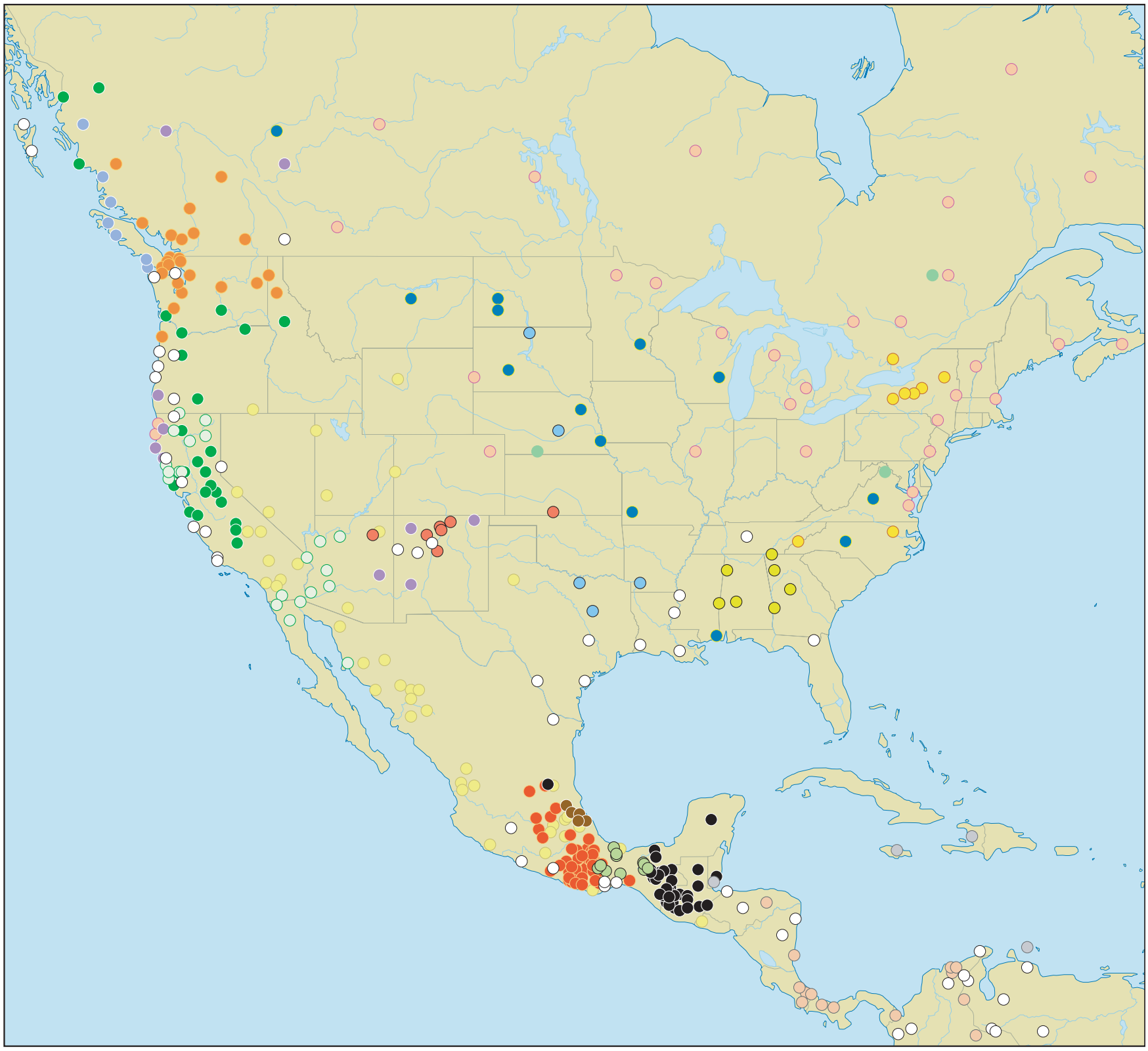
① Ainu
Altaic
Andamanese
Austro-Asiatic
Austronesian
② Burushaski
Creoles and Pidgins
Dravidian
Hmong-Mien
Indo-European
③ Japanese
④ Korean
⑤ Kusunda
⑥ Nahali
Nivkh
Sign Languages
Sino-Tibetan
Tai-Kadai
Uralic
Yeniseian

Pacific



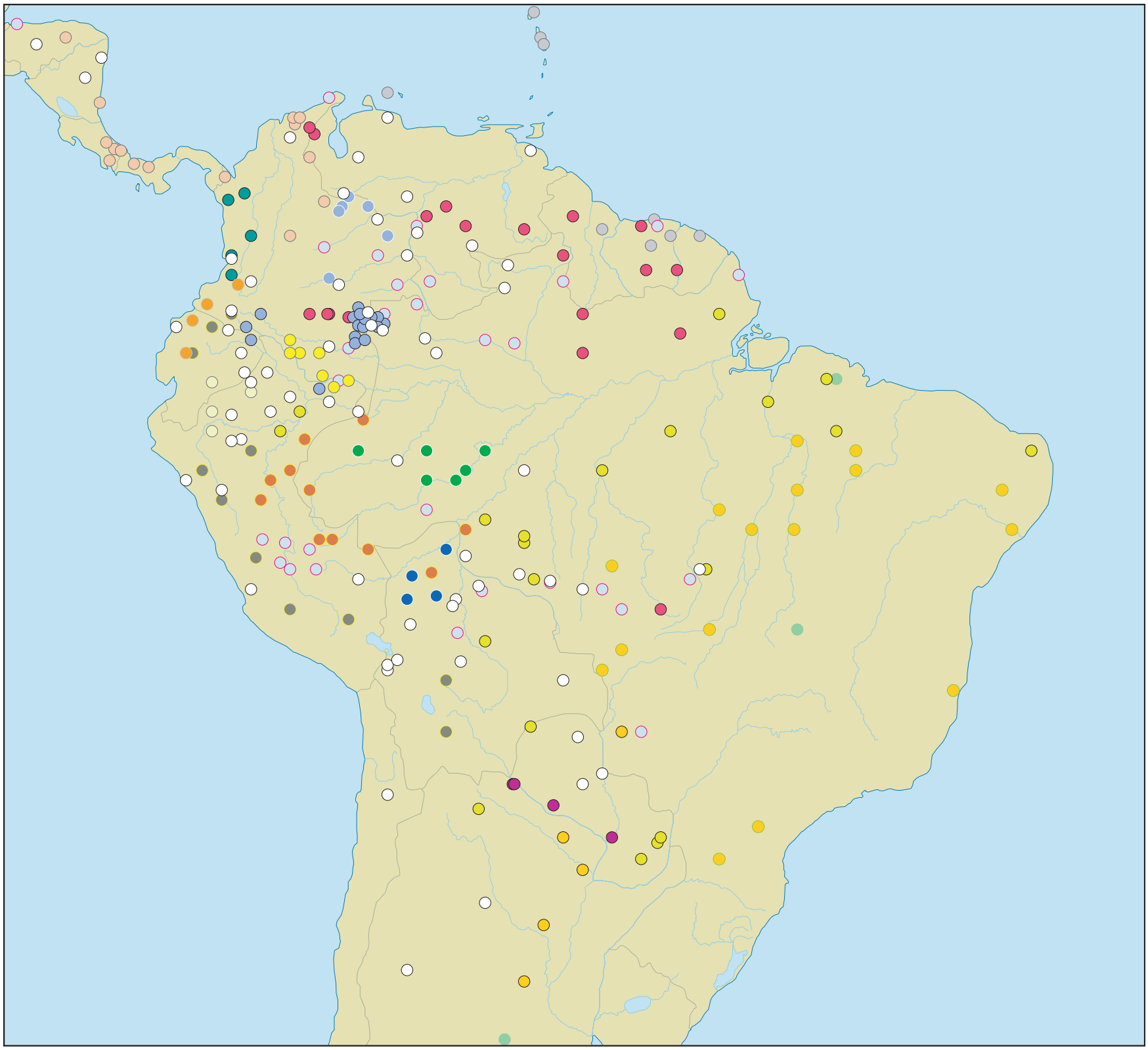
- Australian
- Austronesian
- Baining-Taulil
- Border
- Creoles and Pidgins
- East Bougainville
- Eleman
- Lake Plains
- Lower Sepik-Ramu
- Marind
- Morehead and Upper Maro Rivers
- Sepik
- Sko
- Solomons East Papuan
- Teberan-Pawaian
- Trans-New Guinea
- Torricelli
- West Papuan
- Other Families and Isolates

North America



- Algic
- Caddoan
- Chibchan
- Creoles and Pidgins
- Hokan
- Iroquoian
- Kiowa-Tanoan
- Mayan
- Mixe-Zoque
- Muskogean
- Na-Dene
- Oto-Manguean
- Penutian
- Salishan
- Siouan
- Sign Languages
- Totonacan
- Uto-Aztecan
- Wakashan
- Other Families and Isolates

South America



- Arawakan
- Arauan
- Barbacoan
- Cariban
- Chibchan
- Choco
- Creoles and Pidgins
- Guahiban
- Guaicuruan
- Huitotoan
- Jivaroan
- Macro-Ge
- Mataco
- Panoan
- Quechuan
- Sign Languages
- Tacanan
- Tucanoan
- Tupian
- Other Families and Isolates

Genealogical Language List

BY MATTHEW S. DRYER

This section gives a list of the 2560 languages that appear on at least one map in the atlas, organized according to their genealogical classification. The classification here does not include many intermediate groupings that appear in *Ethnologue* (Grimes 2000) or Ruhlen (1987), restricting itself to only two levels in the majority of cases, that of family, the highest level widely accepted by specialists, and genus, explained presently. In some cases, we provide an intermediate level, that of subfamily, where the subfamily is well known.

The notion "genus" is explained in Dryer (1989a). It is intended as a level of classification which is comparable across the world, so that a genus in one family is intended to be comparable in time depth to genera in other parts of the world. The choice of term is intended to match the general idea of genus in biological classification, where a genus is a set of species that are clearly closely related to each other (and where words in everyday language often correspond to genera rather than species). In the genealogical classification of languages, a genus is a group of languages whose relatedness is fairly obvious without systematic comparative analysis, and which even the most conservative "splitter" would accept. Genealogical groups deeper than a genus are often less obvious and in the absence of detailed comparative work are often not universally accepted. If there is evidence of time depth of groups, the genus would not have a time depth greater than 3500 or 4000 years. A genus may have a time depth much less than this, but if the time of the split of one group of languages from other languages in the family appears to be greater than 4000 years, then this constitutes a reason to say that this group of languages is a separate genus. The standard subfamilies of Indo-European (e.g. Germanic, Slavic, Celtic, etc.) are fairly clear examples of genera, although Celtic is perhaps a clearer example than Germanic or Slavic, both of which have a time depth considerably less than 3500 years. The decisions as to which groups to treat as genera here are best described as my own educated guesses. In many instances they are based on conversations I have had with specialists. However, in the absence of a tradition within the field of attempting to identify groups of comparable time depth in different parts of the world, they should not be considered more than educated guesses. Specialists who think the choices of genera here are mistaken are encouraged to let me know.

The classification here generally follows the classification given in Grimes (2000), the fourteenth edition of *Ethnologue*. Footnotes below indicate differences between the *WALS* classification and that in *Ethnologue*. The classification deviates most strongly from that in *Ethnologue* in the classification of the languages of New Guinea. The *Ethnologue* classification is a standard one, and is similar to that in Ruhlen (1987), both being largely based on the classification in Wurm (1975, 1982). However, most specialists are now sceptical of various components of the Wurm classification. The classification used here is considerably more conservative than the Wurm classification, positing a much larger number of families, and reflects suggestions of William Foley, though I accept responsibility for any misinterpretation of his ideas.

Each family, subfamily, and genus in the classification below is assigned a numeric code consisting of one, two, or three numbers. The first number identifies the family, the second number if there are three numbers identifies the subfamily within the family, and the last number identifies the genus. If a family consists of a single genus (as many do in this classification), then the family is assigned a single number and it is assumed that this single number identifies the genus as well. Names of genera are given in italics. Language isolates are a special case of this: they are families that consist not only of only one genus, but also of only one language. There are cases here of what are generally considered language isolates but for which more than one variety is included in the *WALS* languages. When these varieties are all mutually intelligible, then the language can be considered a language isolate, though since we do not attempt to systematically distinguish languages and dialects (see §3.2 of the Introduction), one cannot tell which families with more than one language here are language isolates in this sense. Examples of language isolates in this sense are Basque and Yukaghir.

The families are organized here in a roughly geographical fashion, starting in southern Africa and ending, except for the last two groups, in southern South America. The order is similar in spirit, but different in detail, from that in Ruhlen (1987). Within each family, the subfamilies and genera are organized alphabetically, and within each genus, the languages are organized alphabetically. The last two groups here are not genealogical groups. The first of these is Creoles and Pidgins, whose nature falls outside standard genealogical classification. The last is Sign Languages. Here the notion of genealogical classification makes sense, but cross-linguistic work on sign languages is sufficiently new that it would be premature to attempt a genealogical classification of them.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
1. KHOISAN[1]						
1.1. Central Khoisan						
//Ani	ani	Botswana	18°55S	21°55E	/Anda	HNH
Deti	det	Botswana	20°30S	24°30E	Deti	DET
Khoekhoe	kho	Namibia	25°30S	18°E	Nama	NAQ
Korana	krn	South Africa	29°30S	20°30E	Korana	KQZ
Kxoe	kxo	Angola, Namibia	17°30S	22°30E	Kxoe	XUU
Nharo	nro	Botswana, Namibia	22°20S	20°30E	Naro	NHR
1.2. Northern Khoisan						
Ju|'hoan	juh	Angola, Namibia, Botswana	19°S	21°E	Ju|'hoan	KTZ
!Xun (Ekoka)	xun	Namibia, Angola	19°40S	18°E	Kung-Ekoka	KNW
1.3. Sandawe						
Sandawe	sdw	Tanzania	5°S	35°E	Sandawe	SBR

[1] The status of Khoisan as a family is controversial. *Ethnologue* also includes Hadza, which is treated here as a language isolate.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
1.4. Southern Khoisan						
/Xam	xam	South Africa	31°S	21°E	/Xam	XAM
!Xóõ	xoo	Botswana	24°S	21°30E	!Xoo	NMN
2. HADZA[2]						
Hadza	had	Tanzania	3°45S	35°10E	Hadza	HTS
3. NIGER-CONGO						
3.1. Adamawa-Ubangian						
Baka (in Cameroon)	bak	Cameroon	2°35N	13°35E	Baka	BKC
Barambu	bbu	Democratic Republic of Congo	3°30N	27°E	Barambu	BRM
Day	day	Chad	8°45N	17°50E	Day	DAI
Dongo	dgo	Democratic Republic of Congo	3°N	30°E	Dongo	DOO
Doyayo	doy	Cameroon	8°40N	13°05E	Doyayo	DOW
Gbaya Kara	gbk	Central African Republic, Cameroon	6°N	15°E	Gbaya (Northwest)	GYA
Gbeya Bossangoa	gbb	Central African Republic	6°40N	17°30E	Gbaya-Bossangoa	GBP
Gula Iro	gir	Chad	10°15N	19°25E	Gula Iro	GLJ
Koh (Lakka)	klk	Chad, Cameroon	8°N	15°30E	Kuo	KHO
Kosop	ksp	Chad	9°55N	15°55E	dialect of Kim	dKIA
Linda	lnd	Central African Republic	6°30N	20°45E	dialect of Banda-Bambari	dLIY
Lua	lua	Chad	9°45N	17°45E	Niellim	NIE
Ma	ma	Democratic Republic of Congo	3°30N	28°E	Ma	MSJ
Mba	mba	Democratic Republic of Congo	1°N	25°E	Mba	MFC
Mbodomo	mdo	Central African Republic	4°30N	15°30E	dialect of Gbaya (Southwest)	dMDO
Mbum	mbm	Cameroon, Chad, Central African Republic	7°45N	13°10E	Mbum	MDD
Mondunga	mga	Democratic Republic of Congo	2°15N	21°30E	Ndunga	NDT
Mumuye	mum	Nigeria	9°N	11°40E	Mumuye	MUL
Mundang	mdg	Chad	9°40N	14°30E	Mundang	MUA
Mündü	mdu	Sudan, Democratic Republic of Congo	4°30N	30°E	Mündü	MUH
Ndogo	ndg	Sudan	7°45N	27°E	Ndogo	NDZ
Ngbaka	ngb	Democratic Republic of Congo	3°30N	20°E	Ngbaka	NGA
Ngbandi	ndi	Democratic Republic of Congo	3°45N	22°E	2 varieties: Ngbandi (Northern)/ Ngbandi (Southern)	NGB/NBW
Nzakara	nza	Central African Republic, Democratic Republic of Congo	5°N	23°E	Nzakara	NZK
Samba Leko	sle	Cameroon, Nigeria	8°30N	12°30E	Samba Leko	NDI
Sango	san	Central African Republic	5°N	18°E	Sango	SAJ
Tupuri	tpr	Cameroon, Chad	10°10N	14°45E	Tupuri	TUI
Yag Dii	ygd	Cameroon	8°N	14°E	Dii	DUR
Yakoma	ykm	Central African Republic	4°25N	18°25E	Yakoma	YKY
Zande	zan	Democratic Republic of Congo, Central African Republic, Sudan	4°N	26°E	Zande	ZAN
3.2. Atlantic						
3.2.1. Northern Atlantic						
Balanta	blz	Senegal, Guinea-Bissau	12°15N	15°20W	2 varieties: Balanta-Ganja/ Balanta-Kentohe	BJY/BLE
Basari	bsr	Senegal, Guinea	12°40N	13°W	Basari	BSC
Biafada	bfd	Guinea-Bissau	11°25N	15°10W	Biafada	BIF
Diola-Fogny	dio	Senegal, Gambia	13°N	16°W	Jola-Fogny	DYO
Ful (Liptako)	fli	Burkina Faso	13°30N	0°30E	dialect of Fulfulde (Northeastern Burkina Faso)	dFUH
Fula (Burkina Faso)	fbf	Burkina Faso	14°N	0°W	Fulfulde (Northeastern Burkina Faso)	FUH
Fula (Cameroonian)	fua	Cameroon	9°10N	13°30E	Fulfulde (Adamawa)	FUB
Fula (Guinean)	fgu	Guinea	11°30N	12°30W	Fuuta Jalon	FUF
Fula (Mauritanian)	fma	Mauritania	16°30N	13°45W	dialect of Pulaar	dFUC
Fula (Nigerian)	fni	Nigeria	8°10N	10°30E	Fulfulde (Nigerian)	FUV
Fula (Senegal)	fus	Senegal	15°N	14°W	Pulaar	FUC
Fulani (Gombe)	fgo	Nigeria	10°15N	11°15E	dialect of Fulfulde (Adamawa)	dFUB
Fulfulde (Maasina)	fum	Mali	15°N	5°W	Fulfulde (Maasina)	FUL
Konyagi	kgi	Guinea, Mali	12°30N	13°15W	Wamei	COU
Manjaku	mjk	Guinea-Bissau	12°N	16°15W	Mandjak	MFV
Mankanya	mkn	Guinea-Bissau	12°05N	15°55W	Mankanya	MAN
Ndut	ndt	Senegal	14°55N	16°55W	Ndut	NDV

[2] Treated by *Ethnologue* as belonging to Khoisan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Noon	noo	Senegal	14°50N	16°50W	Noon	SNF
Palor	plr	Senegal	14°50N	16°45W	Palor	FAP
Wolof	wlf	Senegal, Gambia	15°15N	16°W	Wolof	WOL
3.2.2. Southern Atlantic						
Baga Sitemu	bgs	Guinea	10°30N	14°30W	Baga Sitemu	BSP
Gola	gol	Liberia, Sierra Leone	7°15N	10°50W	Gola	GOL
Kisi	kis	Guinea, Liberia, Sierra Leone	8°50N	10°10W	2 varieties: Kisi (Southern)/ Kissi (Northern)	KSS/KQS
Kisi (Southern)	kss	Liberia, Guinea	8°30N	10°15W	Kisi (Southern)	KSS
Temne	tne	Sierra Leone	8°40N	13°05W	Themne	TEJ
3.3. Benue-Congo						
3.3.1. Bantoid						
Aghem	agh	Cameroon	6°40N	10°E	Aghem	AGQ
Aka	aka	Central African Republic, Congo	2°40N	16°15E	Yaka (in Central African Republic)	AXK
Akwa	awk	Nigeria	10°05N	6°20E	dialect of Pongu	dPON
Babole	bbl	Congo	1°05N	17°15E	Babole	BVX
Babungo	bab	Cameroon	6°07N	10°25E	Vengo	BAV
Bafia	bfi	Cameroon	5°N	11°10E	Bafia	KSF
Bafut	baf	Cameroon	6°15N	9°55E	Bafut	BFD
Bakueri	bku	Cameroon	4°15N	9°15E	Mokpwe	BRI
Bakundu	bkn	Cameroon	4°50N	9°20E	Bakundu-Balue	BDU
Bamun	bmn	Cameroon	5°30N	10°55E	Bamun	BAX
Bankon	bnk	Cameroon	4°25N	9°35E	Abo	ABB
Basaá	bas	Cameroon	3°55N	10°30E	Basaa	BAA
Beembe	bee	Congo	3°55S	14°05E	Beembe	BEJ
Bemba	bem	Democratic Republic of Congo	10°S	28°15E	Bemba	BMY
Bena-Lulua	blu	Democratic Republic of Congo	6°S	21°E	dialect of Luba-Kasai	dLUB
Benga	bga	Equatorial Guinea, Gabon	1°10N	9°25E	Benga	BEN
Bira	bia	Democratic Republic of Congo	1°N	28°45E	2 varieties: Bera/Bila	BRF/BIP
Bobangi	bob	Democratic Republic of Congo	1°20S	17°20E	Bangi	BNI
Bolia	bol	Democratic Republic of Congo	1°15S	18°20E	Bolia	BLI
Bubi	bub	Equatorial Guinea	3°30N	8°40E	Bube	BVB
Bujeba	buj	Cameroon, Equatorial Guinea	3°N	10°10E	Ngumba	NMG
Bukusu	buk	Kenya	0°45N	34°40E	Bukusu	BUL
Bulu	buw	Cameroon	3°N	11°E	Bulu	BUM
Bushoong	bsh	Democratic Republic of Congo	4°30S	21°30E	Bushoong	BUF
Chaga	cga	Tanzania	3°S	37°E	Chagga	KAF
Chichewa	cic	Malawi, Zambia, Mozambique, Zimbabwe	14°S	34°E	Nyanja	NYJ
CiLuba	cil	Democratic Republic of Congo	6°S	22°E	Luba-Kasai	LUB
Comorian	com	Comoros Islands, Mayotte	12°S	44°E	Comorian	SWB
Dabida	dbd	Kenya	4°S	38°40E	Taita	DAV
Ding	dng	Democratic Republic of Congo	4°20S	19°30E	Ding	DIZ
Doko	dok	Democratic Republic of Congo	3°N	23°05E	dialect of Ngombe (Bantu)	dNGC
Duala	dua	Cameroon	4°N	9°25E	Duala	DOU
Duma	dma	Gabon	0°57S	13°E	Duma	DMA
Ejagham	eja	Nigeria, Cameroon	5°25N	8°40E	Ejagham	ETU
Ekoti	eko	Mozambique	16°30S	39°30E	Koti	EKO
Enya	eny	Democratic Republic of Congo	0°30N	25°15E	Enya	GEY
Ewondo	ewo	Cameroon	4°N	12°E	Ewondo	EWO
Fe'fe'	fef	Cameroon	5°15N	10°10E	Fe'fe'	FMP
Fiote	fio	Democratic Republic of Congo	5°30S	14°E	dialect of Kongo	dKON
Gunu	gun	Cameroon	4°35N	11°15E	Nugunu	YAS
Gusii	gus	Kenya	0°45S	34°50E	Gusii	GUZ
Haya	hya	Tanzania	2°S	31°30E	Haya	HAY
Hehe	heh	Tanzania	8°S	36°E	Hehe	HEH
Hemba	hem	Democratic Republic of Congo	6°15S	27°10E	Hemba	HEM
Herero	her	Namibia, Botswana	20°35S	19°E	Herero	HER
Holoholo	hol	Tanzania	5°10S	29°55E	Holoholo	HOO
Hunde	hnd	Democratic Republic of Congo	1°10S	28°50E	Hunde	HKE
Ifumu	ifm	Congo	3°40S	15°20E	Teke (South Central)	IFM
Ila	ila	Zambia, Zimbabwe	15°35S	26°30E	Ila	ILB
Jarawa (in Nigeria)	jar	Nigeria	9°30N	10°30E	Jarawa	JAR
Kaguru	kgr	Tanzania	6°20S	37°E	Kagulu	KKI
Kalanga	klz	Botswana, Zimbabwe	20°30S	27°30E	Kalanga	KCK
Kamba	kba	Kenya	1°30S	38°E	Kamba	KIK
Kami	kmi	Tanzania	6°45S	38°E	Kami	KCU
Kanyok	kyo	Democratic Republic of Congo	7°20S	23°35E	Kanyok	KNY
Karanga	krg	Zimbabwe	20°S	31°E	dialect of Shona	dSHD
Kenyang	kyg	Cameroon	5°40N	9°30E	Kenyang	KEN
Kete	kte	Democratic Republic of Congo	7°S	22°50E	Kete	KCV
Kikuyu	kik	Kenya	0°45S	36°45E	Gikuyu	KIU

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Kiluba	kil	Democratic Republic of Congo	8°S	26°E	Luba-Katanga	LUH
Kinga	kga	Tanzania	9°S	34°E	Kinga	KIX
Kinyarwanda	kin	Rwanda	2°S	30°E	Rwanda	RUA
Kom	kou	Cameroon	6°15N	10°20E	Kom	BKM
Kongo	kon	Democratic Republic of Congo	5°S	15°E	Kongo	KON
Kwangali	kwn	Namibia, Angola	18°S	19°30E	Kwangali	KWN
Lamba	lmb	Zambia, Democratic Republic of Congo	13°S	28°E	Lamba	LAB
Lamnso'	lns	Cameroon	6°10N	10°45E	Lamnso'	NSO
Langi	lgi	Tanzania	4°30S	36°E	Rangi	LAG
Lebeo	leb	Democratic Republic of Congo	1°45N	25°E	Ngelima	AGH
Lega	leg	Democratic Republic of Congo	2°50S	27°10E	2 varieties: Lega-Shabunda/ Lega-Mwenga	LEA/LGM
Lingala	lin	Democratic Republic of Congo	2°N	18°30E	Lingala	LIN
Londo	ldo	Cameroon	5°N	9°10E	Balundu-Bima	NGO
Lozi	loz	Zambia	17°50S	26°E	Lozi	LOZ
Lucazi	luc	Angola	14°S	20°E	Luchazi	LCH
Luganda	lda	Uganda	0°30N	32°10E	Ganda	LAP
Lunda	lbu	Democratic Republic of Congo	10°40S	24°E	Lunda	LVN
Luvale	luv	Angola	12°S	22°E	Luvale	LUE
Luyia	luy	Kenya	0°25N	34°30E	Luyia	LUY
Makaa	mkz	Cameroon	3°25N	12°15E	Makaa	MCP
Makonde	mkd	Tanzania, Mozambique	11°S	40°E	Makonde	KDE
Makua	mua	Mozambique	15°S	38°40E	Makhuwa	VMW
Mambila	mla	Nigeria, Cameroon	6°45N	11°30E	2 varieties: Mambila (Cameroon)/ Mambila (Nigeria)	MYA/MZK
Mambwe	mmw	Zambia, Tanzania	9°05S	31°10E	Mambwe-Lungu	MGR
Matuumbi	mtb	Tanzania	8°30S	39°E	Matumbi	MGW
Mbalanhu	mhu	Namibia	17°30S	15°E	Mbalanhu	LNB
Mbe'	mbz	Cameroon	6°17N	11°05E	Mbe'	MTK
Mbere	mbe	Gabon, Congo	0°30S	14°E	Mbere	MDT
Mbili	mbi	Cameroon	6°07N	10°12E	Bambili	BAW
Mbole	mbl	Democratic Republic of Congo	0°40S	24°40E	Mbole	MDQ
Mituku	mit	Democratic Republic of Congo	1°40S	25°30E	Mituku	ZMQ
Mongo	mgo	Democratic Republic of Congo	0°N	21°E	Mongo (Nkundu)	MOM
Mpongwe	mpo	Gabon	1°15S	9°30E	dialect of Myene	dMYE
Mundani	mud	Cameroon	5°40N	9°55E	Mundani	MUN
Mungaka	mgk	Cameroon	5°55N	10°E	Mungaka	MHK
Mwera	mwe	Tanzania	9°S	39°E	Mwera	MWE
Nande	nde	Democratic Republic of Congo	0°05S	29°10E	Nandi (Bantu)	NNB
Ndebele (in South Africa)	ndb	South Africa	25°20S	29°E	Ndebele (in South Africa)	NEL
Ndonga	ndo	Namibia, Angola	18°S	17°E	Ndonga	NDG
Ndumu	ndu	Gabon	1°40S	13°35E	Ndumu	NMD
Ngemba	nge	Cameroon	6°12N	10°E	Ngemba	NGE
Ngombe	nbe	Democratic Republic of Congo	3°N	23°E	Ngombe (Bantu)	NGC
Ngoni	ngo	Tanzania, Mozambique	11°S	36°E	Ngoni	NGU
Nkem	nke	Nigeria	6°35N	8°35E	dialect of Nkem-Nkum	dISI
Nkore-Kiga	nko	Uganda	0°55S	29°50E	Chiga	CHG
Noni	non	Cameroon	6°25N	10°35E	Noone	NHU
Nsenga	nse	Zambia	14°30S	30°50E	Nsenga	NSE
Ntomba	nto	Democratic Republic of Congo	2°S	18°20E	Ntomba	NTO
Nyambo	nbo	Tanzania	1°S	31°20E	Nyambo	NYM
Nyamwezi	nym	Tanzania	5°S	33°E	Nyamwezi	NYZ
Nyanga	nng	Democratic Republic of Congo	1°15S	28°10E	Nyanga	NYA
Nyiha	nyh	Tanzania, Zambia	10°S	33°E	Nyiha	NIH
Oku	oku	Cameroon	6°15N	10°30E	Oku	OKU
Pangwa	png	Tanzania	10°S	34°45E	Pangwa	PBR
Pare	pre	Tanzania	4°S	36°30E	Asu	ASA
Pogoro	pog	Tanzania	8°45S	36°30E	Pogolo	POY
Rimi	rim	Tanzania	5°S	34°40E	Nyaturu	RIM
Ronga	rga	Mozambique	26°S	32°30E	Ronga	RON
Rundi	rnd	Burundi	3°30S	30°E	Rundi	RUD
Runyankore	rny	Uganda	0°30S	30°30E	Nyankore	NYN
Runyoro-Rutooro	rru	Uganda	1°30N	31°20E	Nyoro	NYR
Sangu	sgu	Gabon	1°30S	11°50E	Sangu	SNQ
Sena	sen	Mozambique	18°S	35°15E	Sena	SEH
Sengele	sgl	Democratic Republic of Congo	1°50S	17°20E	Sengele	SZG
Sesotho	ses	South Africa, Lesotho	28°S	27°E	Sotho (Southern)	SSO
Shambala	shm	Tanzania	4°45S	39°10E	Shambala	KSB
Shona	shn	Zimbabwe	18°S	31°E	Shona	SHD
Songe	sge	Democratic Republic of Congo	6°50S	22°E	Songe	SOP
Sotho (Northern)	stn	South Africa	24°S	29°E	Sotho (Northern)	SRT
Subiya	sub	Namibia	17°45S	24°55E	Subiya	SBS
Sukuma	skm	Tanzania	2°50S	32°E	Sukuma	SUA
Swahili	swa	Tanzania	6°30S	39°E	Swahili	SWA
Swati	swt	Swaziland, South Africa	26°30S	31°E	Swati	SWZ

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Tabwa	tbw	Democratic Republic of Congo, Zambia	8°S	29°30E	Taabwa	TAP
Talinga	tal	Uganda, Democratic Republic of Congo	0°25N	29°40E	Talinga-Bwisi	TLJ
Teke (Southern)	tks	Congo	2°20S	14°30E	Teke (Southern)	KKW
Tetela	tet	Democratic Republic of Congo	4°S	24°E	Tetela	TEL
Tikar	tik	Cameroon	5°50N	11°40E	Tikar	TIK
Tiv	tiv	Nigeria	7°30N	9°E	Tiv	TIV
Tonga (in Zambia)	toz	Zambia	17°S	27°E	Tonga	TOI
Tsogo	tgo	Gabon	1°30S	11°20E	Tsogo	TSV
Tsonga	tsn	South Africa, Mozambique	24°S	32°E	Tsonga	TSO
Tswana	tsw	Botswana, South Africa	24°S	26°E	Tswana	TSW
Tuki	tki	Cameroon	4°35N	11°30E	Tuki	BAG
Tunen	tnn	Cameroon	4°45N	10°40E	Tunen	BAZ
UMbundu	umb	Angola	12°30S	15°E	Umbundu	MNF
Venda	ven	South Africa, Zimbabwe	22°S	30°E	Venda	VEN
Xhosa	xho	South Africa	32°S	27°E	Xhosa	XOS
Yaka	yak	Democratic Republic of Congo, Angola	7°S	17°30E	Yaka (in Democratic Republic of Congo)	YAF
Yamba	yba	Cameroon	6°25N	11°05E	Yamba	YAM
Yansi	yns	Democratic Republic of Congo	3°50S	18°E	Yansi	YNS
Yao (in Malawi)	yao	Malawi, Mozambique	14°30S	35°30E	Yao	YAO
Yemba	yem	Cameroon	5°25N	10°05E	Yemba	BAN
Yeyi	yey	Botswana, Namibia	20°S	23°30E	Yeyi	YEY
Zulu	zul	South Africa	30°S	30°E	Zulu	ZUU
Zulu (Northern)	zno	South Africa	28°S	31°E	dialect of Zulu	dZUU
Zulu (Southern)	zso	South Africa	31°S	30°E	dialect of Zulu	dZUU
3.3.2. Cross River						
Efik	efi	Nigeria	4°55N	8°30E	Efik	EFK
Gokana	gok	Nigeria	4°35N	7°20E	Gokana	GKN
Ibibio	ibi	Nigeria	5°15N	7°50E	Ibibio	IBB
Kana	kan	Nigeria	4°45N	7°25E	Khana	KEH
Kohumono	koh	Nigeria	6°N	8°07E	Kohumono	BCS
Obolo	obo	Nigeria	4°35N	7°50E	Obolo	ANN
Ogbia	ogb	Nigeria	4°40N	6°15E	Ogbia	OGB
Ogbronuagum	obg	Nigeria	4°45N	6°52E	Ogbronuagum	OGU
3.3.3. Defoid						
Isekiri	ise	Nigeria	5°40N	5°30E	Isekiri	ITS
Yoruba	yor	Nigeria, Benin	8°N	4°20E	Yoruba	YOR
3.3.4. Edoid						
Bini	bni	Nigeria	6°25N	5°50E	Edo	EDO
Degema	deg	Nigeria	4°45N	6°45E	Degema	DEG
Engenni	egn	Nigeria	5°10N	6°23E	Engenni	ENN
Etsako	ets	Nigeria	7°15N	6°30E	Yekhee	ETS
Ghotuo	gho	Nigeria	7°05N	5°45E	Ghotuo	AAA
Isoko	iso	Nigeria	5°30N	6°15E	Isoko	ISO
Ora (in Nigeria)	ora	Nigeria	7°05N	5°55E	dialect of Emai-Iuleha-Ora	dEMA
Urhobo	urh	Nigeria	5°35N	6°E	Urhobo	URH
3.3.5. Idomoid						
Idoma	ido	Nigeria	6°55N	7°30E	Idoma	IDO
Igede	ige	Nigeria	7°N	8°10E	Igede	IGE
3.3.6. Igboid						
Igbo	igb	Nigeria	6°N	7°20E	Igbo	IGR
Izi	izi	Nigeria	6°20N	8°E	dialect of Izi-Ezaa-Ikwo-Mgbo	dIZI
3.3.7. Kainji						
Amo	amo	Nigeria	10°20N	8°40E	Amo	AMO
Duka	duk	Nigeria	11°10N	5°05E	Hun-Saare	DUD
3.3.8. Nupoid						
Ebira	ebi	Nigeria	8°10N	7°E	Ebira	IGB
Gade	gad	Nigeria	8°40N	7°25E	Gade	GED
Gwari	gwa	Nigeria	9°30N	7°E	2 varieties: Gbagyi/Gbari	GBR/GBY
Nupe	nup	Nigeria	9°10N	5°25E	Nupe-Nupe-Tako	NUP
3.3.9. Platoid						
Birom	bir	Nigeria	9°40N	8°50E	Berom	BOM
Fyem	fye	Nigeria	9°35N	9°20E	Fyam	PYM

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Gworok	gwo	Nigeria	9°55N	8°10E	dialect of Tyap	dKCG
Jukun	juk	Nigeria	6°55N	10°25E	Jukun Takum	JBU
Kagoma	kgm	Nigeria	9°15N	8°10E	Kagoma	KDM
Kpan	kpa	Nigeria	7°35N	10°10E	Kpan	KPK
Mada (in Nigeria)	mdz	Nigeria	8°45N	8°15E	Mada	MDA
Tarok	tok	Nigeria	9°N	10°05E	Tarok	YER
3.4. Dogon						
Dogon	dog	Mali, Burkina Faso	14°20N	3°30W	Dogon	DOG
Dogon (Toro So)	dts	Mali, Burkina Faso	14°30N	3°20W	dialect of Dogon	dDOG
3.5. Gur						
Bariba	brb	Benin, Nigeria	10°N	2°30E	Bariba	BBA
Bimoba	bmb	Ghana	10°30N	0°E	Bimoba	BIM
Buli (in Ghana)	buy	Ghana	10°30N	1°15W	Buli	BWU
Dagaare	dga	Ghana, Burkina Faso	10°30N	2°40W	3 varieties: Dagaare (Southern)/. . .	DGA/. . .
Dagbani	dgb	Ghana	9°35N	0°30W	Dagbani	DAG
Dyimini	dym	Côte d'Ivoire	8°25N	4°25W	Senoufou (Djimini)	DYI
Gurenne	grn	Ghana	10°50N	0°55W	Frafra	GUR
Gurma	grm	Burkina Faso	12°15N	1°E	Gourmanchéma	GUX
Gurma (Togo)	gut	Burkina Faso	10°45N	0°40E	dialect of Gourmanchéma	dGUX
Kabiyé	kby	Togo	9°40N	1°10E	Kabiyé	KBP
Kasem	ksm	Burkina Faso	11°15N	1°15W	Kasem	KAS
Kirma	kir	Burkina Faso	10°15N	4°50W	Cerma	GOT
Ko (Winye)	kow	Burkina Faso	11°52N	2°55W	Winyé	KST
Konkomba	kkb	Ghana, Togo	10°N	0°05E	Konkomba	KOS
Konni	kni	Ghana	10°15N	1°35W	Konni	KMA
Koromfe	kfe	Burkina Faso, Mali	14°15N	0°55W	Koromfé	KFZ
Lobi	lob	Burkina Faso, Côte d'Ivoire	10°30N	3°15W	Lobi	LOB
Lyele	lye	Burkina Faso	12°30N	2°40W	Lyélé	LEE
Mampruli	mmp	Ghana, Togo	10°20N	0°40W	Mampruli	MAW
Mooré	moo	Burkina Faso	12°50N	1°15W	Mòoré	MHM
Nafaanra	naf	Ghana, Côte d'Ivoire	8°N	2°35W	Nafaanra	NFR
Nateni	ntn	Benin	10°30N	1°10E	Nateni	NTM
Nuni (Northern)	nnn	Burkina Faso	12°10N	3°W	Nuni (Northern)	NUV
Senadi	snd	Côte d'Ivoire	9°30N	6°15W	Senoufou (Cebaara)	SEF
Sisaala	ssa	Ghana	10°50N	2°40W	Sisaala (Western)	SSL
Supyire	sup	Mali	11°30N	5°35W	Senoufo (Supyire)	SPP
Tampulma	tmp	Ghana	10°25N	0°35W	Tampulma	TAM
Tem	tem	Togo, Benin, Ghana	8°40N	0°30E	Tem	KDH
Tenyer	tny	Burkina Faso	10°40N	4°30W	dialect of Karaboro (Western)	dKZA
Toussian	tou	Burkina Faso	11°N	4°40W	2 varieties: Toussian (Northern)/ Toussian (Southern)	TSP/WIB
Toussian (Win)	tow	Burkina Faso	10°50N	4°40W	Toussian (Southern)	WIB
Vagla	vag	Ghana	9°25N	2°25W	Vagla	VAG
3.6. Ijoid						
Defaka	def	Nigeria	4°35N	7°30E	Defaka	AFN
Ijo (Kolokuma)	ijo	Nigeria	4°55N	5°40E	Izon	IJC
3.7. Kordofanian						
Jomang	jom	Sudan	10°35N	30°30E	Talodi	TLO
Katla	ktl	Sudan	11°50N	29°20E	Katla	KCR
Lafofa	laf	Sudan	10°15N	31°15E	Lafofa	LAF
Masakin	msk	Sudan	10°40N	30°E	Ngile	MAS
Moro	mro	Sudan	11°N	30°10E	Moro	MOR
Orig	ori	Sudan	12°10N	30°50E	dialect of Tegali	dRAS
Otoro	otr	Sudan	11°10N	30°30E	Otoro	OTR
Rashad	ras	Sudan	11°50N	31°03E	dialect of Tegali	dRAS
3.8. Kru						
Aizi	aiz	Côte d'Ivoire	5°15N	4°30W	3 varieties: Aizi (Aproumu)/. . .	AHP/. . .
Bété	bet	Côte d'Ivoire	6°15N	6°15W	3 varieties: Bété (Daloa)/. . .	BEV/. . .
Godié	gdi	Côte d'Ivoire	5°25N	5°50W	Godié	GOD
Grebo	grb	Liberia	5°N	8°W	5 varieties: Grebo (Barclayville)/. . .	GRY/. . .
Guere	gue	Côte d'Ivoire	6°40N	7°45W	2 varieties: Wè Southern/ Wè Western	GXX/WEC
Klao	kla	Liberia	4°45N	8°45W	Klao	KLU
Seme	sme	Burkina Faso	11°N	4°55W	Siamou	SIF

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Vata	vat	Côte d'Ivoire	5°50N	5°25W	dialect of Dida (Lakota)	dDIC
Wobe	wob	Côte d'Ivoire	7°25N	7°20W	Wè Northern	WOB
3.9. Kwa						
Abidji	abd	Côte d'Ivoire	5°40N	4°35W	Abidji	ABI
Adioukrou	adi	Côte d'Ivoire	5°25N	4°35W	Adioukrou	ADJ
Akan	akn	Ghana	6°30N	1°15W	Akan	TWS
Alladian	ald	Côte d'Ivoire	5°10N	4°20W	Alladian	ALD
Anufo	anu	Ghana, Benin, Togo	10°05N	0°15E	Anufo	CKO
Anyi	ayi	Côte d'Ivoire	7°N	3°30W	Anyin	ANY
Avatime	avt	Ghana	6°50N	0°25E	Avatime	AVA
Baule	ble	Côte d'Ivoire	7°N	5°W	Baoulé	BCI
Chumburung	cum	Ghana	7°45N	0°15E	Chumburung	NCU
Ega	ega	Côte d'Ivoire	5°30N	5°30W	Ega	DIE
Ewe	ewe	Ghana, Togo	6°20N	0°25E	Éwé	EWE
Ewe (Anglo)	ewa	Ghana	6°N	1°E	dialect of Éwé	dEWE
Fongbe	fon	Benin	6°25N	2°10E	Fon-Gbe	FOA
Gã	ga	Ghana	5°40N	0°10W	Ga	GAC
Kposo	kpo	Togo	7°30N	0°50E	Akposos	KPO
Lelemi	llm	Ghana	7°20N	0°30E	Lelemi	LEF
Nkonya	nkn	Ghana	7°10N	0°15E	Nkonya	NKO
Wémênugbé	wmn	Benin	6°40N	2°30E	Weme-Gbe	WEM
3.10. Mande						
3.10.1. Eastern Mande						
Beng	bnq	Côte d'Ivoire	7°50N	4°15W	Beng	NHB
Bisa	bis	Burkina Faso, Ghana	11°30N	0°30W	Bissa	BIB
Boko	bok	Benin, Nigeria	10°30N	3°30E	Boko	BQC
Busa	bus	Nigeria	9°40N	4°E	Busa	BQP
Dan	dan	Côte d'Ivoire, Guinea, Liberia	7°30N	8°W	Dan	DAF
Dan (Blowo)	dbw	Côte d'Ivoire	7°20N	8°10W	dialect of Dan	dDAF
Guro	gro	Côte d'Ivoire	7°N	6°W	Guro	GOA
Mano	man	Liberia, Guinea	7°N	9°W	Mano	MEV
3.10.2. Western Mande						
Bambara	bam	Mali	12°30N	7°30W	Bambara	BRA
Bandi	bnd	Liberia	8°05N	10°15W	Bandi	GBA
Bobo Fing	bbf	Burkina Faso	11°50N	4°25W	2 varieties: Bobo Madaré (Northern)/ Bobo Madaré (Southern)	BBO/BWQ
Bozo (Tigemaxo)	boz	Mali	15°N	4°W	Bozo (Tiéyaxo)	BOZ
Dyula	dyu	Burkina Faso, Mali, Côte d'Ivoire	9°50N	4°40W	Jula	DYU
Jeli	jel	Côte d'Ivoire	9°30N	5°40W	—	—
Koranko	kko	Sierra Leone, Guinea	9°20N	11°15W	Kuranko	KHA
Kpelle	kpe	Liberia	7°N	10°W	Kpelle	KPE
Loko	lok	Sierra Leone	9°20N	12°05W	Loko	LOK
Loma	lma	Liberia, Guinea	8°N	9°30W	Loma	LOM
Mandinka	mdk	Senegal, Guinea, Mali	13°N	15°40W	Mandinka	MNK
Mandinka (Gambian)	mkg	Gambia	13°25N	16°W	dialect of Mandinka	dMNK
Maninka	mnk	Guinea, Sierra Leone	10°30N	9°30W	Maninka (Kankan)	MNI
Maninka (Western)	maw	Mali	13°N	11°W	Maninkakan (Western)	MLQ
Mauka	mka	Côte d'Ivoire	8°20N	7°30W	Mahou	MXX
Mende	mde	Sierra Leone	7°50N	11°30W	Mende	MFY
Soninke	snn	Mali, Senegal, Côte d'Ivoire, Burkina Faso	15°N	10°30W	Soninke	SNN
Susu	sus	Guinea	10°N	13°W	Susu	SUD
Vai	vai	Liberia, Sierra Leone	6°55N	11°15W	Vai	VAI
Xasonga	xas	Mali, Senegal	14°15N	10°30W	Xasonga	KAO
4. LAAL[3]						
Laal	laa	Chad	10°N	17°40E	Laal	GDM
5. KADUGLI[4]						
Kadugli	kad	Sudan	11°N	29°40E	dialect of Katcha-Kadugli-Miri	dKAT
Katcha	ktc	Sudan	10°50N	29°40E	dialect of Katcha-Kadugli-Miri	dKAT
Krongo	kro	Sudan	10°30N	30°E	Krongo	KGO

[3] Treated by *Ethnologue* as Niger-Congo.
[4] Treated by *Ethnologue* as a branch of Kordofanian within Niger-Congo. Bender (1997) treats it as a branch of Nilo-Saharan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
6. NILO-SAHARAN[5]						
6.1. Berta						
Berta	ber	Ethiopia, Sudan	10°20N	34°40E	Berta	WTI
6.2. Central Sudanic						
6.2.1. Bongo-Bagirmi						
Bagirmi	bag	Chad	11°40N	16°E	Bagirmi	BMI
Bagiro	bgr	Democratic Republic of Congo, Central African Republic	4°20N	20°35E	Furu	FUU
Baka (in Sudan)	bka	Sudan	4°50N	29°15E	Baka	BDH
Binga	big	Central African Republic, Sudan	8°N	25°E	dialect of Yulu	dYUL
Birri	brz	Central African Republic	5°30N	25°10E	Birri	BVQ
Bongo	bgo	Sudan	7°30N	28°30E	Bongo	BOT
Gula (in Central African Republic)	gul	Central African Republic	9°30N	22°30E	Gula	GLU
Jur Mödö	jmo	Sudan	6°N	30°E	Jur Modo	BEX
Kara (in Central African Republic)	kar	Central African Republic	10°N	23°E	Kara	KAH
Kenga	ken	Chad	12°N	18°E	Kenga	KYQ
Mbay	mby	Chad	8°15N	17°30E	Mbay	MYB
Ngambay	ngm	Chad, Cameroon	8°40N	16°E	Ngambay	SBA
Yulu	yul	Sudan	8°30N	25°15E	Yulu	YUL
6.2.2. Kresh						
Aja	aja	Sudan, Central African Republic	7°20N	25°40E	Aja	AJA
Kresh	kre	Sudan	8°30N	24°30E	Gbaya	KRS
6.2.3. Lendu						
Lendu	ldu	Democratic Republic of Congo	2°N	30°30E	Lendu	LED
Ngiti	nti	Democratic Republic of Congo	1°20N	30°15E	Ngiti	NIY
6.2.4. Mangbetu						
Mangbetu	mbt	Democratic Republic of Congo	2°30N	26°30E	Mangbetu	MDJ
6.2.5. Mangbutu-Efe						
Lese	les	Democratic Republic of Congo	2°N	29°E	Lese	LES
Mamvu	mmv	Democratic Republic of Congo	3°15N	29°E	Mamvu	MDI
6.2.6. Moru-Ma'di						
Avokaya	avo	Democratic Republic of Congo, Sudan	4°40N	30°E	Avokaya	AVU
Logoti	lgt	Democratic Republic of Congo	3°30N	30°E	Logo	LOG
Lugbara	lug	Uganda, Democratic Republic of Congo	3°05N	30°55E	Lugbara	LUG
Ma'di	mad	Uganda, Sudan	3°15N	31°30E	Ma'di	MHI
Moru	mou	Sudan	4°45N	29°45E	Moru	MGD
6.3. Eastern Sudanic						
6.3.1. Daju						
Daju (Dar Fur)	ddf	Sudan	12°15N	25°15E	Daju (Dar Fur)	DAJ
Shatt	sht	Sudan	10°50N	30°E	Shatt	SHJ
Sila	sil	Chad	12°20N	21°45E	Daju (Dar Sila)	DAU
6.3.2. Eastern Jebel						
Ingessana	igs	Sudan	11°30N	34°E	Gaam	TBI
6.3.3. Kuliak						
Ik	ik	Uganda	3°45N	34°10E	Ik	IKX
Nyangi	nyg	Uganda	3°25N	33°35E	Nyang'i	NYP
So	so	Uganda	2°35N	34°45E	Soo	TEU
6.3.4. Nara						
Nara (in Ethiopia)	nar	Eritrea	15°05N	37°35E	Nara	NRB

[5] The exact membership of Nilo-Saharan is somewhat uncertain. The membership of Songhay in Nilo-Saharan is questioned by some, and it might be treated as a separate family.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
6.3.5. Nilotic						
Acholi	acl	Uganda, Sudan	3°N	32°40E	Acholi	ACO
Anywa	any	Ethiopia	8°N	33°30E	Anuak	ANU
Bari	bar	Sudan	5°N	31°40E	Bari	BFA
Datooga	dat	Tanzania	4°20S	35°20E	Datooga	TCC
Dinka	din	Sudan	8°30N	28°E	5 varieties: Dinka (Northeastern)/…	DIP/...
Jur Luwo	jlu	Sudan	8°N	28°E	Luwo	LWO
Karimojong	kmj	Uganda	3°N	34°10E	Karamojong	KDJ
Kukú	kuk	Sudan	5°50N	31°40E	dialect of Bari	dBFA
Lango	lan	Uganda	2°10N	33°E	Lango	LAJ
Luo	luo	Kenya, Tanzania	0°30S	34°45E	Luo	LUO
Maasai	maa	Kenya, Tanzania	3°S	36°E	Maasai	MET
Nandi	nan	Kenya	0°15N	35°E	dialect of Kalenjin	dKLN
Nuer	nue	Sudan, Ethiopia	8°20N	32°E	Nuer	NUS
Päri	par	Sudan	5°N	32°30E	Päri	LKR
Pokot	pkt	Kenya, Uganda	1°30N	35°30E	Pökoot	PKO
Sebei	seb	Uganda	1°20N	34°35E	Kupsabiny	KPZ
Shilluk	shl	Sudan	9°40N	31°45E	Shilluk	SHK
Teso	tes	Uganda	1°50N	33°50E	Teso	TEO
Turkana	tna	Kenya, Uganda	3°N	35°30E	Turkana	TUV
6.3.6. Nubian						
Midob	mid	Sudan	15°05N	26°15E	Midob	MEI
Nobiin	nob	Sudan, Egypt	21°N	31°E	Nobiin	FIA
Nubian (Dongolese)	nbd	Sudan	18°15N	30°45E	dialect of Kenuzi-Dongola	dKNC
Nubian (Hill)	nbh	Sudan	12°N	31°E	7 varieties: Ghulfan/. . .	GHL/. . .
Nubian (Kunuz)	nku	Egypt	23°N	33°E	dialect of Kenuzi-Dongola	dKNC
6.3.7. Nyimang						
Nyimang	nyi	Sudan	12°10N	29°20E	Ama	NYI
6.3.8. Surmic						
Baale	blj	Sudan	6°30N	34°30E	dialect of Kacipo-Balesi	dKOE
Chai	cai	Ethiopia	5°20N	35°20E	dialect of Suri	dSUQ
Didinga	did	Sudan	4°30N	33°30E	Didinga	DID
Koegu	koe	Ethiopia	7°N	36°05E	Kwegu	YID
Majang	maj	Ethiopia	6°45N	35°E	Majang	MPE
Me'en	mee	Ethiopia	5°N	37°E	Me'en	MYM
Murle	mrl	Sudan	6°30N	33°30E	Murle	MUR
Mursi	mur	Ethiopia	5°35N	36°05E	Mursi	MUZ
Tennet	ten	Sudan	4°25N	32°30E	Tennet	TEX
Tirmaga	trm	Ethiopia	5°45N	35°15E	dialect of Suri	dSUQ
6.3.9. Taman						
Miisiirii	mii	Sudan, Chad	13°30N	22°E	dialect of Tama	dTMA
Sungor	sug	Chad	13°45N	21°30E	Assangori	SUN
Tama	tma	Sudan, Chad	14°30N	22°E	Tama	TMA
6.3.10. Temein						
Temein	tmn	Sudan	11°55N	29°25E	Temein	TEQ
6.4. Fur						
Fur	fur	Sudan	13°30N	25°E	Fur	FUR
6.5. Gumuz						
Gumuz	gmz	Ethiopia, Sudan	12°30N	35°50E	Gumuz	GUK
6.6. Koman						
Komo	kom	Sudan, Ethiopia	8°45N	33°45E	Komo	KOM
6.7. Kunama						
Kunama	knm	Ethiopia, Eritrea	14°30N	37°E	Kunama	KUM
6.8. Maban						
Maba	mab	Chad	13°45N	20°50E	Maba	MDE
Masalit	msl	Sudan, Chad	13°20N	22°E	Masalit	MSA
Runga	run	Chad, Central African Republic	10°20N	21°E	Runga	ROU

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
6.9. Saharan						
Kanembu	kbu	Chad, Nigeria, Niger	14°N	15°E	Kanembu	KBL
Kanuri	knr	Nigeria, Niger, Sudan, Chad	12°N	13°E	3 varieties: Kanuri (Central)/. . .	KPH/. . .
Tubu	tbu	Chad, Niger	16°N	15°E	2 varieties: Dazaga/Tedaga	DAK/TUQ
6.10. Shabo						
Shabo	shb	Ethiopia	7°35N	35°25E	Shabo	SBF
6.11. Songhay						
Koyra Chiini	kch	Mali	17°N	3°W	Songhay (Koyra Chiini)	KHQ
Koyraboro Senni	kse	Mali, Niger	16°N	0°E	Songhay (Koyraboro Senni)	SON
Zarma	zar	Niger	13°50N	2°10E	Zarma	DJE
7. AFRO-ASIATIC						
7.1. Berber						
Berber (Ayt Seghrouchen Middle Atlas)	bse	Morocco	33°20N	4°40W	dialect of Tamazight (Central Atlas)	dTZM
Berber (Chaouia)	bch	Algeria	35°N	7°E	Chaouia	SHY
Berber (Figuig)	bfg	Morocco	32°30N	1°30W	—	—
Berber (Middle Atlas)	bma	Morocco	33°N	5°W	Tamazight (Central Atlas)	TZM
Berber (Mzab)	bmz	Algeria	32°30N	3°30E	Tumzabt	MZB
Berber (Rif)	brf	Morocco, Algeria	34°30N	4°W	Tarifit	RIF
Berber (Siwa)	bsi	Egypt	29°10N	25°30E	Siwi	SIZ
Berber (Wargla)	bou	Algeria	31°55N	5°20E	Tagargrent	OUA
Kabyle	kbl	Algeria	36°30N	5°E	Kabyle	KYL
Tashlhiyt	tas	Morocco	31°N	5°W	Tachelhit	SHI
Tuareg (Ahaggar)	tug	Algeria, Libya, Niger	23°N	6°E	Tamahaq (Tahaggart)	THV
Tuareg (Air)	tai	Niger	19°N	9°E	dialect of Tamahaq (Tahaggart)	dTHV
Tuareg (Ghat)	tgh	Libya	25°N	10°10E	dialect of Tamahaq (Tahaggart)	dTHV
Zenaga	zen	Mauritania	17°20N	16°W	Zenaga	ZEN
7.2. Chadic						
7.2.1. Biu-Mandara						
Bana	ban	Cameroon	10°25N	13°35E	Bana	BCW
Buduma	bud	Chad, Cameroon, Nigeria	13°30N	14°30E	Buduma	BDM
Bura-Pabir	bpa	Nigeria	10°30N	12°15E	Bura-Pabir	BUR
Daba	dab	Cameroon	10°10N	13°45E	Daba	DAB
Ga'anda	gnd	Nigeria	10°20N	12°35E	Ga'anda	GAA
Gidar	gdr	Cameroon	10°N	14°E	Gidar	GID
Giziga	giz	Cameroon	10°20N	14°10E	2 varieties: Giziga (North)/ Giziga (South)	GIS/GIZ
Gude	gud	Nigeria, Cameroon	10°25N	13°25E	Gude	GDE
Guduf	gdf	Nigeria, Cameroon	11°14N	13°48E	Guduf	GDF
Hdi	hdi	Cameroon, Nigeria	11°52N	13°43E	Hdi	TUR
Kilba	klb	Nigeria	10°20N	13°10E	Huba	KIR
Kotoko	ktk	Central African Republic	11°20N	15°20E	Afade	AAL
Lagwan	lag	Cameroon	11°30N	14°50E	Lagwan	KOT
Lamang	lmg	Nigeria	11°15N	13°35E	Lamang	HIA
Mada (in Cameroon)	mda	Cameroon	10°50N	14°10E	Mada	MXU
Malgwa	mlg	Nigeria	11°30N	13°45E	dialect of Wandala	dMFI
Margi	mrg	Nigeria	11°N	13°E	Marghi (Central)	MAR
Mbara	mbr	Chad	11°N	15°25E	Mbara	MPK
Mofu-Gudur	mof	Cameroon	10°30N	14°E	Mofu-Gudur	MIF
Musgu	mgu	Cameroon, Chad	10°50N	14°55E	Musgu	MUG
Podoko	pod	Cameroon	11°N	14°05E	Parkwa	PBI
Tera	ter	Nigeria	11°N	11°50E	Tera	TER
Uldeme	uld	Cameroon	10°57N	14°07E	Wuzlam	UDL
7.2.2. East Chadic						
Bidiya	bid	Chad	11°55N	18°45E	Bidiyo	BID
Dangaléat (Western)	dnw	Chad	12°10N	18°20E	dialect of Dangaléat	dDAA
Kera	ker	Chad	9°50N	15°05E	Kera	KER
Lele	lel	Chad	9°05N	15°35E	Lele	LLN
Migama	mig	Chad	12°10N	19°50E	Migaama	MMY
Mokilko	mko	Chad	11°55N	18°05E	Mukulu	MOZ
Tumak	tmk	Chad	9°30N	17°20E	Tumak	TMC

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
7.2.3. Masa						
Lamé	lam	Chad, Cameroon	9°N	14°30E	Pévé	LME
Masa	mas	Chad, Cameroon	10°30N	15°30E	Masana	MCN
7.2.4. West Chadic						
Angas	anc	Nigeria	9°30N	9°30E	Ngas	ANC
Bade	bad	Nigeria	12°50N	10°30E	Bade	BDE
Bole	blq	Nigeria	11°20N	11°15E	Bole	BOL
Goemai	goe	Nigeria	8°40N	9°45E	Goemai	ANK
Hausa	hau	Nigeria, Niger	12°N	7°E	Hausa	HUA
Kanakuru	knk	Nigeria	10°N	12°E	Dera	KNA
Kwami	kwm	Nigeria	10°25N	11°E	Kwaami	KSQ
Miya	miy	Nigeria	10°55N	9°40E	Miya	MKF
Mupun	mup	Nigeria	9°30N	8°50E	Mwaghavul	SUR
Ngizim	ngz	Nigeria	12°05N	10°55E	Ngizim	NGI
Pa'a	paa	Nigeria	11°N	9°15E	Pa'a	AFA
Pero	per	Nigeria	9°35N	11°E	Pero	PIP
Ron	ron	Nigeria	9°N	8°45E	Ron	CLA
Tangale	tan	Nigeria	9°45N	11°20E	Tangale	TAN
7.3. Cushitic						
7.3.1. Beja						
Beja	bej	Sudan, Eritrea	18°N	36°E	Bedawi	BEI
7.3.2. Central Cushitic						
Awngi	awn	Ethiopia	10°50N	36°40E	Awngi	AWN
Bilin	bln	Eritrea	15°45N	38°30E	Bilen	BYN
Kemant	kem	Ethiopia	12°40N	37°25E	dialect of Agaw (Western)	dQIM
7.3.3. Eastern Cushitic						
Arbore	abo	Ethiopia	5°N	36°45E	Arbore	ARV
Bayso	bys	Ethiopia	6°15N	37°45E	Baiso	BSW
Burji	brj	Ethiopia	5°30N	37°50E	Burji	BJI
Dhaasanac	dha	Ethiopia	4°40N	36°20E	Daasanach	DSH
Dullay (Gollango)	dug	Ethiopia	5°30N	37°15E	Gawwada	GWD
Oromo (Boraana)	orb	Ethiopia, Kenya	4°30N	38°30E	dialect of Borana-Arsi-Guji	dGAX
Oromo (Harar)	orh	Ethiopia	9°N	42°E	Oromo (Eastern)	HAE
Oromo (Waata)	orw	Kenya	3°20S	39°50E	Sanye	SSN
Oromo (West-Central)	owc	Ethiopia	9°N	37°E	Oromo (West-Central)	GAZ
Qafar	qaf	Ethiopia, Eritrea, Djibouti	12°N	42°E	Afar	AFR
Rendille	ren	Kenya	2°N	37°30E	Rendille	REL
Saho	sao	Eritrea, Ethiopia	14°45N	39°45E	Saho	SSY
Somali	som	Somalia	3°N	45°E	Somali	SOM
7.3.4. Southern Cushitic						
Alagwa	agw	Tanzania	5°30S	35°45E	Wasi	WBJ
Burunge	brn	Tanzania	5°20S	36°E	Burunge	BDS
Dahalo	dah	Kenya	2°20S	40°30E	Dahalo	DAL
Gorowa	gor	Tanzania	4°30S	36°30E	Gorowa	GOW
Iraqw	irq	Tanzania	4°S	35°30E	Iraqw	IRK
Mbugu	mbg	Tanzania	4°20S	38°10E	Mbugu	MHD
7.4. Egyptian-Coptic						
Coptic	cop	Egypt	26°N	32°E	Coptic	COP
7.5. Omotic						
Aari	aar	Ethiopia	6°N	36°35E	Aari	AIZ
Dime	dim	Ethiopia	6°10N	36°15E	Dime	DIM
Dizi	diz	Ethiopia	6°10N	36°30E	Dizi	MDX
Gamo	gam	Ethiopia	6°40N	37°15E	dialect of Gamo-Gofa-Dauro	dGMO
Gimira	gim	Ethiopia	7°N	35°45E	Bench	BCQ
Hamer	hmr	Ethiopia	5°N	36°30E	Hamer-Banna	AMF
Kefa	kef	Ethiopia	7°15N	36°15E	Kaficho	KBR
Koyra	kjr	Ethiopia	5°50N	37°50E	Koorete	KQY
Kullo	kul	Ethiopia	6°45N	37°05E	dialect of Gamo-Gofa-Dawro	dGMO
Maale	mle	Ethiopia	6°N	36°55E	Male	MDY
Moca	moc	Ethiopia	7°35N	35°30E	Shakacho	MOY
Shinassha	ssh	Ethiopia	10°25N	36°10E	Boro (Omotic)	BWO

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Wolaytta	wly	Ethiopia	6°50N	37°45E	Wolaytta	WBC
Yemsa	yms	Ethiopia	7°50N	37°20E	Yemsa	JNJ
Zayse	zay	Ethiopia	5°55N	37°20E	Zayse-Zergulla	ZAY

7.6. Semitic

Amharic	amh	Ethiopia	10°N	38°E	Amharic	AMH
Arabic (Abbéché Chad)	abb	Chad	13°50N	20°50E	dialect of Arabic (Chadian Spoken) (= Arabic (Shuwa))	dSHU
Arabic (Bahrain)	abh	Bahrain	26°N	50°30E	Arabic (Baharna Spoken)	AFH
Arabic (Bani-Hassan)	ahs	Mauritania	20°N	10°W	Hassaniyya	MEY
Arabic (Beirut)	abe	Lebanon	33°55N	35°30E	dialect of Arabic (North Levantine Spoken)	dAPC
Arabic (Borno Nigerian)	arn	Nigeria	8°35N	12°25E	dialect of Arabic (Shuwa) (= Arabic (Chadian Spoken))	dSHU
Arabic (Eastern Libyan)	ael	Libya	32°N	22°E	dialect of Arabic (Libyan Spoken)	dAYL
Arabic (Egyptian)	aeg	Egypt	30°N	31°E	Arabic (Egyptian Spoken)	ARZ
Arabic (Gulf)	arg	Saudi Arabia, Kuwait, Bahrain, Qatar, Iran, Iraq, Oman	26°N	49°E	Arabic (Gulf Spoken)	AFB
Arabic (Hijazi)	arh	Saudi Arabia	26°N	38°E	Arabic (Hijazi Spoken)	ACW
Arabic (Iraqi)	arq	Iraq	33°N	44°E	Arabic (Mesopotamian Spoken)	ACM
Arabic (Kormakiti)	ako	Cyprus	35°15N	33°30E	dialect of Arabic (Cyproit Spoken)	dACY
Arabic (Lebanese)	arl	Lebanon	34°N	36°E	dialect of Arabic (North Levantine Spoken)	dAPC
Arabic (Modern Standard)	ams	Saudi Arabia, Kuwait, Bahrain, Qatar, United Arab Emirates, Oman, Yemen, Palestinian West Bank and Gaza, Israel, Iraq, Syria, Lebanon, Jordan, Egypt, Libya, Tunisia, Algeria, Morocco, Somalia, Sudan, Eritrea, Comoros Islands	25°N	42°E	Arabic (Standard)	ABV
Arabic (Moroccan)	amr	Morocco	34°N	6°W	Arabic (Moroccan Spoken)	ARY
Arabic (Negev)	arv	Israel	30°30N	34°45E	dialect of Arabic (South Levantine Spoken)	dAJP
Arabic (North Levantine Spoken)	anl	Lebanon	33°45N	36°E	Arabic (North Levantine Spoken)	APC
Arabic (Palestinian)	apa	Palestinian West Bank and Gaza	32°N	35°15E	Arabic (South Levantine Spoken)	AJP
Arabic (San'ani)	ars	Yemen	16°N	44°E	Arabic (Sanaani)	AYN
Arabic (Syrian)	asy	Syria	34°N	38°E	dialect of Arabic (North Levantine Spoken)	dAPC
Arabic (Tunisian)	atu	Tunisia	36°45N	10°15E	Arabic (Tunisian Spoken)	AEB
Bathari	bth	Oman	18°N	56°E	dialect of Mehri	dMHR
Chaha	chh	Ethiopia	8°N	37°30E	dialect of Gurage (West)	dGUY
Chaldean (Modern)	cld	Iraq	36°N	43°E	Chaldean Neo-Aramaic	CLD
Harari	hrr	Ethiopia	9°15N	42°10E	Harari	HAR
Harsusi	hrs	Oman	20°N	56°30E	Harsusi	HSS
Hebrew (Modern Ashenazic)	hba	Israel	31°45N	35°10E	dialect of Hebrew	dHBR
Hebrew (Modern)	heb	Israel	31°30N	34°50E	Hebrew	HBR
Jibbali	jib	Oman	17°30N	55°E	Jibbali	SHV
Maltese	mlt	Malta	35°55N	14°25E	Maltese	MLS
Mandaic (Modern)	mma	Iran, Iraq	31°N	48°30E	Mandaic	MID
Mehri	meh	Yemen	18°N	53°E	Mehri	MHR
Muher	muh	Ethiopia	8°10N	38°10E	dialect of Gurage (West)	dGUY
Neo-Aramaic (Amadiya)	naa	Iraq	37°N	43°E	dialect of Chaldean (Neo-Aramaic)	dCLD
Neo-Aramaic (Arbel Jewish)	naj	Israel	31°45N	35°E	Lishanid Noshan	AIJ
Neo-Aramaic (Assyrian)	nsy	Syria	36°N	41°E	Neo-Aramaic (Assyrian)	AII
Neo-Aramaic (Persian Azerbaijan)	nap	Iran, Israel	38°N	47°E	Lishán Didán	TRG
Soddo	sod	Ethiopia	8°30N	38°30E	Gurage (Soddo)	GRU
Soqotri	soq	Yemen	12°30N	54°E	Soqotri	SQT
Tigré	tgr	Eritrea	16°30N	38°30E	Tigré	TIE
Tigré (Beni Amer)	tbe	Eritrea	15°30N	38°E	dialect of Tigré	dTIE
Tigrinya	tig	Ethiopia, Eritrea	14°30N	38°30E	Tigrigna	TGN

8. BASQUE

Basque	bsq	Spain, France	43°N	3°W	Basque	BSQ
Basque (Basaburua and Imoz)	bqi	Spain	43°N	1°40W	dialect of Basque	dBSQ
Basque (Bidasoa Valley)	bqb	Spain	43°N	3°50W	dialect of Basque	dBSQ
Basque (Gernica)	bqg	Spain	43°20N	3°W	dialect of Basque	dBSQ

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Basque (Hondarribia)	bqh	Spain	42°30N	3°W	dialect of Basque	dBSQ
Basque (Lekeitio)	bql	Spain	43°N	2°30W	dialect of Basque	dBSQ
Basque (Northern High Navarrese)	bqn	Spain	43°20N	2°30W	dialect of Basque	dBSQ
Basque (Oñati)	bqo	Spain	42°30N	3°30W	dialect of Basque	dBSQ
Basque (Roncalese)	bqr	Spain	43°N	2°W	dialect of Basque	dBSQ
Basque (Sakana)	bqs	Spain	42°30N	2°W	dialect of Basque	dBSQ
Basque (Souletin)	bso	France	43°30N	1°30W	Basque (Souletin)	BSZ
Basque (Zeberio)	bqz	Spain	42°N	3°W	dialect of Basque	dBSQ

9. INDO-EUROPEAN

9.1. Albanian

Albanian	alb	Albania, Serbia and Montenegro	41°N	20°E	2 varieties: Albanian (Tosk)/ Albanian (Gheg)	ALS

9.2. Armenian

Armenian (Eastern)	arm	Armenia	40°N	45°E	dialect of Armenian	dARM
Armenian (Iranian)	arz	Iran	32°40N	51°40E	dialect of Armenian	dARM
Armenian (Western)	arw	Turkey	38°30N	43°30E	dialect of Armenian	dARM

9.3. Baltic

Latvian	lat	Latvia	57°N	24°E	Latvian	LAT
Lithuanian	lit	Lithuania	55°N	24°E	Lithuanian	LIT

9.4. Celtic

Breton	bre	France	48°N	3°W	Breton	BRT
Cornish	crn	United Kingdom	50°20N	5°W	Cornish	CRN
Gaelic (Scots)	gae	United Kingdom	57°N	4°W	Gaelic (Scots)	GLS
Irish	iri	Ireland	53°N	8°W	Gaelic (Irish)	GLI
Irish (Donegal)	ird	Ireland	55°N	8°W	dialect of Gaelic (Irish)	dGLI
Irish (Munster)	irm	Ireland	52°30N	9°W	dialect of Gaelic (Irish)	dGLI
Manx	mnx	United Kingdom	54°10N	4°30W	Manx	MJD
Welsh	wel	United Kingdom	52°N	3°W	Welsh	WLS

9.5. Germanic

Afrikaans	afr	South Africa	31°S	22°E	Afrikaans	AFK
Alsatian	alt	France	48°30N	7°30E	dialect of Alemannisch	dGSW
Danish	dsh	Denmark	56°N	10°E	Danish	DNS
Dutch	dut	Netherlands	52°30N	6°E	Dutch	DUT
Dutch (Brabantic)	dbr	Belgium	50°45N	4°30E	dialect of Dutch	dDUT
Dutch (Limburg)	dli	Belgium	51°N	5°30E	dialect of Dutch	dDUT
Dutch (Zeeuws)	duz	Netherlands, Belgium	51°30N	3°45E	dialect of Vlaams	dVLA
English	eng	United Kingdom, Ireland	52°N	0°W	English	ENG
Faroese	far	Denmark	62°N	7°W	Faroese	FAE
Frisian	fri	Netherlands, Germany	53°N	5°E	3 varieties: Frisian (Eastern)/. . .	FRS/. . .
Frisian (Eastern)	fea	Germany	53°35N	7°30E	Frisian (Eastern)	FRS
Frisian (North)	fno	Germany	54°30N	9°E	Frisian (Northern)	FRR
Frisian (Western)	frw	Netherlands	53°N	6°E	Frisian (Western)	FRI
German	ger	Germany, Austria, Switzerland	52°N	10°E	German (Standard)	GER
German (Appenzell)	gpz	Switzerland	47°20N	9°20E	dialect of Alemannisch	dGSW
German (Bavarian)	gba	Austria, Germany	49°N	11°30E	Bavarian	BAR
German (Berlin)	gbl	Germany	52°30N	13°20E	dialect of German (Standard)	dGER
German (Bern)	gbe	Switzerland	47°N	7°25E	dialect of Allemannisch	dGSW
German (Hannover)	gha	Germany	52°25N	9°40E	dialect of German (Standard)	dGER
German (Mansfeldisch)	gma	Germany	51°35N	11°30E	dialect of German (Standard)	dGER
German (Ostschweiz)	gos	Switzerland	47°25N	9°15E	dialect of Alemannisch	dGSW
German (Ripuarian)	grp	Belgium	51°N	7°E	Kölsch	KOR
German (Thurgau)	gtg	Switzerland	47°35N	9°10E	dialect of Alemannisch	dGSW
German (Thuringian)	gth	Germany	51°N	11°E	dialect of German (Standard)	dGER
German (Timisoara)	gti	Romania	45°45N	21°15E	dialect of German (Standard)	dGER
German (Upper Austrian)	gau	Austria	47°N	14°E	dialect of Bavarian	dBAR
German (Viennese)	gvi	Austria	48°10N	16°20E	dialect of Bavarian	dBAR
German (Westphalian)	gwe	Germany	52°N	7°30E	Westphalien	WEP
German (Zurich)	gzu	Switzerland	47°25N	8°30E	dialect of Alemannisch	dGSW
Icelandic	ice	Iceland	65°N	17°W	Icelandic	ICE
Low German	lge	Germany	53°N	10°E	Saxon (Low)	SXN
Luxemburgeois	lux	Luxembourg	49°50N	6°10E	Luxembourgeois	LUX
Norwegian	nor	Norway	61°N	8°E	Norwegian (Bokmaal)	NRR
Swedish	swe	Sweden, Finland	60°N	15°E	Swedish	SWD

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Swedish (Västerbotten)	swv	Sweden	64°15N	19°45E	dialect of Swedish	dSWD
Yiddish	ydd	Germany, Poland, Belarus, Ukraine, Lithuania	52°N	23°E	2 varieties: Yiddish (Eastern)/ Yiddish (Western)	YDD/YIH
Yiddish (Bessarabian)	ydb	Moldova	47°N	28°30E	dialect of Yiddish (Eastern)	dYDD
Yiddish (Lithuanian)	ylt	Lithuania	55°N	25°E	dialect of Yiddish (Eastern)	dYDD
Yiddish (Lodz)	ydl	Poland	51°45N	19°25E	dialect of Yiddish (Eastern)	dYDD
9.6. Greek						
Greek (Cypriot)	gcy	Cyprus	34°45N	33°E	dialect of Greek	dGRK
Greek (Modern)	grk	Greece	39°N	22°E	Greek	GRK
9.7. Indic						
Assamese	ass	India	26°N	93°E	Assamese	ASM
Awadhi	awd	India, Nepal	27°30N	81°30E	Awadhi	AWD
Bagri	bgi	India	29°35N	74°20E	Bagri	BGQ
Bengali	ben	Bangladesh, India	24°N	90°E	Bengali	BNG
Bengali (Chittagong)	bec	Bangladesh	22°20N	91°50E	Chittagonian	CIT
Bhili	bhi	India	22°N	73°E	Bhili	BHB
Bhojpuri	bho	India	26°N	84°E	Bhojpuri	BHJ
Brokskat	bkt	India, Pakistan	35°20N	76°30E	Brokskat	BKK
Darai	dar	Nepal	24°N	84°E	Darai	DRY
Dhivehi	dhi	Maldives	4°10N	73°30E	Maldivian	SNM
Dogri	dgi	India	33°N	75°E	Dogri	DOJ
Domari	dom	Iran, Israel, Jordan	32°N	35°E	Domari	RMT
Dumaki	dmk	Pakistan	36°10N	74°E	Domaaki	DMK
Gojri	goj	India	32°45N	76°15E	Gujari	GJU
Gujarati	guj	India	23°N	72°E	Gujarati	GJR
Halbi	hlb	India	21°N	81°E	Halbi	HLB
Hindi	hin	India	25°N	77°E	Hindi	HND
Johari	joh	India	30°40N	80°E	dialect of Kumauni	dKFY
Kalami	kal	Pakistan	35°30N	72°30E	Kalami	GWC
Kalasha	klh	Pakistan	35°N	72°E	Kalasha	KLS
Kashmiri	kas	Pakistan, India	34°N	76°E	Kashmiri	KSH
Kati (in Afghanistan)	ktz	Afghanistan, Pakistan	35°30N	70°E	Kati	BSH
Khowar	khw	Pakistan	36°N	72°E	Khowar	KHW
Kokni	kkz	India	20°45N	73°30E	Kukna	KEX
Konkani	kkn	India	15°15N	74°E	Konkani	KNK
Kumauni	kum	India	30°N	80°E	Kumauni	KFY
Lamani	lmn	India	17°N	77°E	Lambadi	LMN
Magahi	mgh	India	23°30N	85°30E	Magahi	MQM
Maithili	mai	India, Nepal	26°N	86°E	Maithili	MKP
Marathi	mhi	India	19°N	76°E	Marathi	MRT
Mawchi	mwc	India	21°20N	73°40E	Mawchi	MKE
Nepali	nep	Nepal	28°N	85°E	Nepali	NEP
Oriya	oya	India	21°N	85°E	Oriya	ORY
Oriya (Kotia)	oko	India	18°20N	83°E	Oriya (Adivasi)	ORT
Panjabi	pan	Pakistan, India	31°N	74°E	Panjabi (Eastern)	PNJ
Prasuni	pra	Afghanistan	35°30N	71°E	Prasuni	PRN
Romani (Ajia Varvara)	rav	Greece	38°N	23°35E	dialect of Romani (Balkan)	dRMN
Romani (Bugurdzi)	rbu	Serbia and Montenegro	42°40N	21°10E	dialect of Romani (Balkan)	dRMN
Romani (Burgenland)	rbg	Austria	47°30N	16°30E	dialect of Romani (Sinte)	dRMO
Romani (Kalderash)	rka	Serbia and Montenegro	45°N	21°E	dialect of Romani (Vlax)	dRMY
Romani (Lovari)	rlo	Hungary, Romania	47°N	21°30E	dialect of Romani (Vlax)	dRMY
Romani (North Russian)	rnr	Russia	60°N	38°E	dialect of Romani (Baltic)	dROM
Romani (Sepecides)	rse	Turkey	38°15N	27°E	—	—
Romani (Welsh)	rwe	United Kingdom	52°N	4°W	Romani (Welsh)	RMW
Savi	sav	Afghanistan	35°N	71°30E	Savi	SDG
Shekhawati	skw	India	28°N	75°30E	dialect of Marwari	dMKD
Shina	sna	Pakistan	36°N	74°E	Shina	SCL
Sindhi	sdh	Pakistan, India	26°N	69°E	Sindhi	SND
Sinhala	snh	Sri Lanka	7°N	80°30E	Sinhala	SNH
Torwali	trw	Pakistan	34°15N	72°E	Torwali	TRW
Urdu	urd	Pakistan	25°N	67°E	Urdu	URD
Vasavi	vas	India	24°30N	71°30E	Vasavi	VAS
Vedda	ved	Sri Lanka	7°N	81°E	Veddah	VED
Waigali	wgl	Afghanistan	35°N	71°15E	Waigali	WBK
9.8. Iranian						
Baluchi	blc	Pakistan, Iran, Afghanistan	28°N	62°E	3 varieties: Balochi (Eastern)/. . .	BGP/. . .
Dari	dri	Afghanistan	35°N	66°E	Farsi (Eastern)	PRS

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Gilaki	gil	Iran	37°20N	50°E	Gilaki	GLK
Hawrami	hwr	Iraq	35°45N	45°30E	Hawrami	HAC
Ishkashmi	ish	Afghanistan, Tajikistan	37°N	72°E	dialect of Sanglechi-Ishkashimi	dSGL
Kabatei	kbt	Iran	36°45N	49°25E	Kabatei	XKP
Koluri	klr	Iran	37°N	48°E	dialect of Shahrudi	dSHM
Kurdish (Central)	krd	Iraq, Iran	36°N	44°E	5 varieties: Behdini/. . .	BDF/. . .
Kurmanji	kji	Turkey	38°N	42°E	Kurmanji	KUR
Luri	lur	Iran	33°30N	48°30E	Luri	LRI
Mazanderani	mzn	Iran	36°30N	52°E	Mazanderani	MZN
Ormuri	orm	Pakistan	32°30N	69°45E	Ormuri	ORU
Ossetic	oss	Russia, Georgia	43°N	44°E	Osetin	OSE
Pashto	psh	Afghanistan, Pakistan	33°N	67°E	4 varieties: Pashto (Central)/. . .	PST/. . .
Persian	prs	Iran	32°N	54°E	Farsi (Western)	PES
Rushan	rsh	Tajikistan, Afghanistan	38°10N	71°30E	dialect of Shugnhi	dSGH
Tajik	taj	Tajikistan	38°40N	70°E	Tajiki	PET
Talysh (Azerbaijan)	taz	Azerbaijan	38°45N	48°40E	dialect of Talysh	dTLY
Talysh (Southern)	tls	Iran	37°N	49°E	dialect of Talysh	dTLY
Tat (Muslim)	tmu	Azerbaijan	41°15N	48°45E	Tat (Muslim)	TTT
Tati (Southern)	tts	Iran	36°N	49°40E	Takestani	TKS
Vafsi	vaf	Iran	34°40N	49°55E	Vafsi	VAF
Wakhi	wak	Afghanistan, Pakistan, Tajikistan	36°30N	72°E	Wakhi	WBL
Yaghnobi	ygn	Tajikistan	39°N	69°E	Yagnobi	YAI
Yazgulyam	yaz	Tajikistan	38°30N	71°30E	Yazgulyam	YAH
Zazaki	zaz	Turkey	39°N	41°E	Dimli	ZZZ

9.9. Romance

Asturian	ast	Spain	43°15N	6°W	Asturian	AUB
Catalan	ctl	Spain	41°45N	2°E	Catalan-Valencian-Balear	CLN
French	fre	France, Switzerland	48°N	2°E	French	FRN
Galician	glc	Spain	43°N	8°W	Galician	GLN
Italian	ita	Italy, Switzerland	43°N	12°E	Italian	ITN
Italian (Bologna)	itb	Italy	44°30N	11°20E	Emiliano-Romagnolo	EML
Italian (Fiorentino)	ifi	Italy	43°45N	11°15E	dialect of Italian	dITN
Italian (Genoa)	itg	Italy	44°25N	8°57E	dialect of Ligurian	dLIJ
Italian (Napolitanian)	itn	Italy	40°55N	14°15E	dialect of Napoletano-Calabrese	dNPL
Italian (Turinese)	itu	Italy	45°N	7°40E	Piemontese	PMS
Ladin	ldn	Italy	46°35N	11°55E	Ladin	LLD
Ladino	lno	Israel, Turkey	40°N	33°E	Ladino	SPJ
Moldavian	mol	Moldova	47°N	29°E	dialect of Romanian	dRUM
Occitan	occ	France	44°N	2°E	4 varieties: Languedocien/. . .	LNC/. . .
Portuguese	por	Portugal	39°N	8°W	Portuguese	POR
Provençal	pro	France	44°N	6°E	Provençal	PRV
Romanian	rom	Romania	46°N	25°E	Romanian	RUM
Romansch	rmc	Switzerland	46°40N	9°10E	Romansch	RHE
Romansch (Scharans)	rsc	Switzerland	46°45N	9°30E	dialect of Romansch	dRHE
Romansch (Surmeiran)	rsm	Switzerland	46°35N	9°45E	dialect of Romansch	dRHE
Romansch (Sursilvan)	rsu	Switzerland	46°40N	8°50E	dialect of Romansch	dRHE
Sardinian	srd	Italy	40°N	9°E	4 varieties: Sardinian (Campidanese)/. . .	SRO/. . .
Spanish	spa	Spain	40°N	4°W	Spanish	SPN
Spanish (Canary Islands)	spc	Spain	28°N	15°35W	dialect of Spanish	dSPN

9.10. Slavic

Belorussian	blr	Belarus	54°N	28°E	Belorussian	RUW
Bosnian	bos	Bosnia-Herzegovina	43°N	18°E	dialect of Serbo-Croatian	dSRC
Bulgarian	bul	Bulgaria	42°30N	25°E	Bulgarian	BLG
Czech	cze	Czech Republic	50°N	15°E	Czech	CZC
Kashubian	ksu	Poland	54°N	18°E	Kashubian	CSB
Macedonian	mcd	Macedonia	41°40N	21°45E	Macedonian	MKJ
Polabian	plb	Germany	51°N	12°50E	Polabian	POX
Polish	pol	Poland	52°N	20°E	Polish	PQL
Russian	rus	Russia	56°N	38°E	Russian	RUS
Serbian-Croatian	scr	Serbia and Montenegro, Croatia, Bosnia-Herzegovina	44°N	19°E	Serbo-Croatian	SRC
Slovak	svk	Slovakia	49°N	20°E	Slovak	SLO
Slovene	slo	Slovenia	46°N	15°E	Slovenian	SLV
Slovincian	svc	Poland	54°N	16°E	dialect of Kashubian	dCSB
Sorbian	srb	Germany	51°30N	14°E	2 varieties: Sorbian (Lower)/ Sorbian (Upper)	WEE/WEN
Sorbian (Lower)	srl	Germany	51°45N	14°20E	Sorbian (Lower)	WEE
Sorbian (Upper)	sou	Germany	51°50N	14°30E	Sorbian (Upper)	WEN
Ukrainian	ukr	Ukraine	49°N	33°E	Ukrainian	UKR

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
10. URALIC						
10.1. Finno-Ugric						
10.1.1. Finnic						
Estonian	est	Estonia	59°N	26°E	Estonian	EST
Finnish	fin	Finland	62°N	25°E	Finnish	FIN
Izhor	izh	Russia	59°N	29°E	Ingrian	IZH
Karelian	krl	Russia	64°N	32°E	Karelian	KRL
Komi-Permyak	kop	Russia	59°30N	54°30E	Komi-Permyak	KOI
Komi-Zyrian	kzy	Russia	65°N	55°E	Komi-Zyrian	KPV
Liv	liv	Latvia	56°50N	24°E	Liv	LIV
Mari (Hill)	mah	Russia	57°N	58°E	Mari (High)	MRJ
Mari (Meadow)	mme	Russia	57°N	48°E	Mari (Low)	MAL
Mordvin (Erzya)	moe	Russia, Kazakhstan	53°N	45°30E	Erzya	MYV
Mordvin (Moksha)	mmo	Russia	54°N	44°E	Moksha	MDF
Saami (Central-South)	scs	Sweden	64°40N	16°45E	Saami (Southern)	LPC
Saami (Kildin)	ski	Russia	67°N	37°E	Saami (Kildin)	LPD
Saami (Northern)	sno	Norway, Finland, Sweden	69°N	24°E	Saami (Northern)	LPR
Udmurt	udm	Russia	57°30N	52°30E	Udmurt	UDM
Veps	vep	Russia	60°N	35°E	Veps	VEP
Votic	vot	Russia	59°30N	30°E	Vod	VOD
Yazva	yzv	Russia	65°N	55°E	dialect of Komi-Zyrian	dKPV
10.1.2. Ugric						
Hungarian	hun	Hungary	47°N	20°E	Hungarian	HNG
Khanty	kty	Russia	65°N	65°E	Khanty	KCA
Mansi	mns	Russia	62°N	62°E	Mansi	MNS
10.2. Samoyedic						
Enets	ene	Russia	67°30N	86°30E	Enets	ENE
Kamass	kms	Russia	52°N	92°E	Kamas	XAS
Nenets	nen	Russia	69°N	72°E	Nenets	YRK
Nenets (Tundra)	ntu	Russia	70°N	76°E	dialect of Nenets	dYRK
Nganasan	nga	Russia	71°N	93°E	Nganasan	NIO
Selkup	skp	Russia	65°N	82°E	Selkup	SAK
11. ALTAIC						
11.1. Mongolic						
Baonan	bao	China	36°N	102°30E	Bonan	PEH
Buriat	but	Mongolia, China, Russia	52°N	108°E	3 varieties: Buriat (Mongolia)/. . .	BXM/. . .
Dagur	dgr	China	48°N	124°E	Daur	DTA
Kalmyk	kmk	Mongolia, Russia, China	44°N	83°E	dialect of Kalmyk-Oirat	dKGZ
Kalmyk (Issyk-Kul)	kiq	Kyrgyzstan	42°25N	78°20E	dialect of Kalmyk-Oirat	dKGZ
Khalkha	kha	Mongolia	47°N	105°E	Mongolian (Halh)	KHK
Mangghuer	mgg	China	36°N	102°E	Tu	MJG
Moghol	mog	Afghanistan	35°N	62°E	Mogholi	MLG
Mongol (Khamnigan)	mkh	China, Mongolia	49°N	117°E	—	—
Oirat	oir	Mongolia, China	37°30N	93°E	dialect of Kalmyk-Oirat	dKGZ
Ordos	ord	China	39°N	109°E	dialect of Mongolian (Peripheral)	dMVF
Santa	sta	China	31°30N	107°30E	Dongxiang	SCE
Shira Yughur	shy	China	38°N	99°E	East Yugur	YUY
11.2. Tungusic						
Even	evn	Russia	68°N	130°E	Even	EVE
Evenki	eve	Russia	56°N	125°E	Evenki	EVN
Jurchen	jur	China	47°N	130°E	Jurchen	JUC
Manchu	mnc	China	49°30N	127°30E	Manchu	MJF
Nanai	nai	Russia	49°30N	137°E	Nanai	GLD
Negidal	neg	Russia	53°N	139°E	Negidal	NEG
Oroch	orc	Russia	50°20N	137°30E	Oroch	OAC
Orok	ork	Russia	49°30N	143°30E	Orok	OAA
Solon	sol	China	48°N	120°E	dialect of Evenki	dEVN
Udihe	udh	Russia	47°40N	136°15E	Udihe	UDE
Ulcha	ulc	Russia	52°15N	140°20E	Ulch	ULC
11.3. Turkic						
Altai (Southern)	aso	Russia	51°N	87°E	Altai (Southern)	ALT
Aynu	ayn	China	39°30N	76°E	Ainu	AIB

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Azari (Iranian)	azi	Iran, Iraq	37°30N	47°E	Azerbaijani (South)	AZB
Azerbaijani	aze	Azerbaijan, Iran	40°30N	48°30E	2 varieties: Azerbaijani (North)/ Azerbaijani (South)	AZE/AZB
Bashkir	bsk	Russia	53°N	58°E	Bashkir	BXK
Chulym	cly	Russia	54°20N	89°40E	Chulym	CHU
Chuvash	chv	Russia	55°30N	47°30E	Chuvash	CJU
Crimean Tatar	cri	Uzbekistan	45°N	34°05E	Crimean Turkish	CRH
Dolgan	dol	Russia	71°15N	98°E	Dolgan	DLG
Gagauz	gag	Moldova	46°20N	28°40E	Gagauz	GAG
Karachay-Balkar	krc	Russia	43°30N	42°E	Karachay-Balkar	KRC
Karaim	krm	Lithuania, Ukraine	54°40N	24°55E	Karaim	KDR
Karakalpak	kkp	Uzbekistan	43°N	60°E	Karakalpak	KAC
Kazakh	kaz	Kazakhstan	50°N	70°E	Kazakh	KAZ
Khakas	khk	Russia	53°N	90°E	Khakas	KJH
Khalaj	khl	Iran	35°N	50°E	Khalaj	KLJ
Kirghiz	kgz	Kyrgyzstan	42°N	75°E	Kirghiz	KDO
Kirghiz (Fu-Yu)	kfy	China	47°45N	124°25E	dialect of Kirghiz	dKDO
Krymchak	kym	Ukraine	45°N	34°15E	dialect of Uzbek (Northern)	dUZB
Kumyk	kuq	Russia	43°N	47°20E	Kumyk	KSK
Noghay	nog	Russia	44°N	46°E	Nogai	NOG
Noghay (Karagash)	nok	Russia	47°30N	47°15E	dialect of Nogai	dNOG
Salar	slr	China	35°N	103°E	Salar	SLR
Saryg Yughur	syg	China	38°45N	99°45E	West Yugur	YBE
Shor	shr	Russia	53°N	88°E	Shor	CJS
Tatar	tvo	Russia	56°N	49°30E	Tatar	TTR
Tatar (Baraba)	ttb	Russia	55°20N	78°20E	dialect of Tatar	dTTR
Tatar (Mishar)	tmi	Russia	55°30N	49°E	dialect of Tatar	dTTR
Tatar-Noghay (Alabugat)	tlb	Russia	47°25N	45°40E	dialect of Nogai	dNOG
Tofa	tof	Russia	53°45N	98°E	Karagas	KIM
Turkic (East-Central Xorasan)	tex	Iran	37°N	59°E	dialect of Khorasani Turkish	dKMZ
Turkic (West Xorasan)	twx	Iran	36°40N	56°E	dialect of Khorasani Turkish	dKMZ
Turkish	tur	Turkey	39°N	35°E	Turkish	TRK
Turkmen	tkm	Turkmenistan	40°N	58°E	Turkmen	TCK
Tuvan	tuv	Russia, Mongolia	52°N	95°E	Tuvin	TUN
Urum	urm	Georgia, Ukraine	47°07N	37°33E	Urum	UUM
Uyghur	uyg	China	40°N	80°E	Uyghur	UIG
Uzbek	uzb	Uzbekistan, Afghanistan	39°N	66°E	2 varieties: Uzbek (Southern)/ Uzbek (Northern)	UZS/UZB
Uzbek (Northern)	uzn	Uzbekistan	40°40N	66°30E	Uzbek (Northern)	UZB
Yakut	ykt	Russia	62°N	130°E	Yakut	UKT
Yurt Tatar	yta	Russia	46°20N	48°30E	dialect of Nogai	dNOG
12. YUKAGHIR						
Yukaghir (Kolyma)	yko	Russia	65°45N	150°50E	Yukaghir (Southern)	YUX
Yukaghir (Tundra)	ytu	Russia	69°N	155°E	Yukaghir (Northern)	YKG
13. YENISEIAN						
Ket	ket	Russia	64°N	87°E	Ket	KET
Kott	ktt	Russia	57°N	94°E	—	—
Yugh	yug	Russia	61°N	90°E	Yugh	YUU
14. CHUKOTKO-KAMCHATKAN						
14.1. Northern Chukotko-Kamchatkan						
Alutor	alu	Russia	61°N	165°E	Alutor	ALR
Chukchi	chk	Russia	67°N	173°W	Chukot	CKT
Kerek	krq	Russia	62°15N	175°E	Kerek	KRK
Koryak	kry	Russia	61°N	167°E	Koryak	KPY
14.2. Southern Chukotko-Kamchatkan						
Itelmen	ite	Russia	57°N	157°30E	Itelmen	ITL
15. NIVKH						
Nivkh	niv	Russia	53°20N	142°E	Gilyak	NIV
Nivkh (South Sakhalin)	nvs	Russia	47°N	142°30E	dialect of Gilyak	dNIV

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
16. AINU						
Ainu	ain	Japan	43°N	143°E	Ainu	AIN
17. JAPANESE						
Japanese	jpn	Japan	37°N	140°E	Japanese	JPN
18. KOREAN						
Korean	kor	South Korea, North Korea	37°30N	128°E	Korean	KKN
19. NORTH-WEST CAUCASIAN[6]						
Abaza	abz	Russia	44°N	42°E	Abaza	ABQ
Abkhaz	abk	Georgia	43°05N	41°E	Abkhaz	ABK
Adyghe (Abzakh)	ady	Russia	45°14N	40°35E	dialect of Adyghe	dADY
Adyghe (Shapsugh)	ash	Russia	45°N	38°45E	dialect of Adyghe	dADY
Adyghe (Temirgoy)	adt	Russia	45°13N	39°42E	dialect of Adyghe	dADY
Kabardian	kab	Russia	43°30N	43°30E	Kabardian	KAB
Ubykh	uby	Russia, Turkey	43°40N	39°40E	Ubykh	UBY
20. NAKH-DAGHESTANIAN[7]						
20.1. Daghestanian						
20.1.1. Avar-Andic-Tsezic						
Akhvakh	axv	Russia	42°20N	46°20E	Akhvakh	AKV
Andi	anx	Russia	42°45N	46°15E	Andi	ANI
Avar	ava	Russia, Azerbaijan	42°30N	46°30E	Avar	AVR
Bagvalal	bgv	Russia	42°35N	46°10E	Bagvalal	KVA
Bezhta	bez	Russia	42°05N	46°10E	Bezhta	KAP
Chamalal	chm	Russia	42°33N	46°03E	Chamalal	CJI
Godoberi	god	Russia	42°40N	46°05E	Ghodoberi	GDO
Hinukh	hnk	Russia	42°10N	46°E	Hinukh	GIN
Hunzib	hzb	Russia	42°10N	46°15E	Hunzib	HUZ
Karata	krt	Russia	42°35N	46°20E	Karata	KPT
Khvarshi	khv	Russia	42°15N	46°10E	Khvarshi	KHV
Tindi	tnd	Russia	42°22N	46°15E	Tindi	TIN
Tsez	tsz	Russia	42°15N	45°45E	Dido	DDO
20.1.2. Lak-Dargwa						
Dargwa	drg	Russia	42°15N	47°25E	Dargwa	DAR
Lak	lak	Russia	42°10N	47°10E	Lak	LBE
20.1.3. Lezgic						
Aghul	agl	Russia	41°45N	47°40E	Aghul	AGX
Archi	arc	Russia	42°N	46°50E	Archi	ARC
Budukh	bdk	Azerbaijan	41°10N	48°25E	Budukh	BDK
Khinalug	khi	Azerbaijan	41°10N	48°05E	Khinalugh	KJJ
Kryz	krz	Azerbaijan	41°05N	48°E	Kryts	KRY
Lezgian	lez	Russia, Azerbaijan	41°40N	47°50E	Lezgi	LEZ
Rutul	rut	Russia, Azerbaijan	41°30N	47°25E	Rutul	RUT
Tabassaran	tbs	Russia	41°50N	47°55E	Tabassaran	TAB
Tsakhur	tsa	Russia, Azerbaijan	41°40N	47°10E	Tsakhur	TKR
Udi	udi	Azerbaijan	41°N	48°E	Udi	UDI
20.2. Nakh						
Chechen	chc	Russia	43°15N	45°50E	Chechen	CJC
Ingush	ing	Russia	43°10N	45°05E	Ingush	INH
Tsova-Tush	ttu	Georgia	42°30N	45°30E	Bats	BBL

[6] Treated by *Ethnologue* as a branch of North Caucasian, which is not recognized here as a family.

[7] The two branches of Nakh-Daghestanian, namely Nakh and Daghestanian, are treated as branches of North Caucasian by *Ethnologue*. While there have been proposals grouping Nakh-Daghestanian and Northwest Caucasian together, the *Ethnologue* classification is unusual in failing to treat Nakh-Daghestanian as a genealogical grouping.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
21. KARTVELIAN						
Georgian	geo	Georgia	42°N	44°E	Georgian	GEO
Laz	laz	Turkey, Georgia	41°30N	41°30E	Laz	LZZ
Mingrelian	mgl	Georgia	42°30N	42°E	Mingrelian	XMF
Svan	sva	Georgia	43°N	42°30E	Svan	SVA
22. BURUSHASKI						
Burushaski	bur	Pakistan	36°30N	74°30E	Burushaski	BSK
23. NAHALI						
Nahali	nah	India	19°45N	77°50E	Nihali	NHL
24. KUSUNDA[8]						
Kusunda	kus	Nepal	28°N	84°40E	Kusunda	KGG
25. DRAVIDIAN						
25.1. Central Dravidian						
Kolami	kol	India	20°N	78°30E	2 varieties: Kolami (Northwestern)/ Kolami (Southeastern)	KFB/NIT
Parji (Dravidian)	prd	India	19°30N	82°30E	Duruwa	PCI
25.2. Northern Dravidian						
Brahui	brh	Pakistan	28°30N	67°E	Brahui	BRH
Kurukh	kur	India	22°50N	85°30E	Kurux	KVN
Malto	mto	India	25°N	87°20E	3 varieties: Paharia (Kumarbhag)/. . .	KMJ/. . .
25.3. South-Central Dravidian						
Gondi	gon	India	19°N	81°E	2 varieties: Gondi (Northern)/ Gondi (Southern)	GON/GGO
Konda	kda	India	19°30N	83°E	dialect of Konda-Dora	dKFC
Koya	koy	India	17°30N	81°20E	Koya	KFF
Kui (in India)	kui	India	20°N	83°30E	Kui	KXU
Kuvi	kuv	India	18°45N	82°40E	Kuvi	KXV
Pengo	pen	India	19°50N	83°50E	Pengo	PEG
Telugu	tel	India	16°N	79°E	Telugu	TCW
25.4. Southern Dravidian						
Badaga	bdg	India	11°40N	76°45E	Badaga	BFQ
Kannada	knd	India	14°N	76°E	Kannada	KJV
Kodava	kod	India	12°10N	76°50E	Kodagu	KFA
Kota	kot	India	11°30N	77°10E	Kota	KFE
Malayalam	mym	India	10°N	76°30E	Malayalam	MJS
Tamil	tml	India, Sri Lanka	11°N	78°30E	Tamil	TCV
Tamil (Spoken)	tsp	India	11°N	78°E	dialect of Tamil	dTCV
Toda	tda	India	11°25N	76°45E	Toda	TCX
Tulu	tul	India	12°45N	75°20E	Tulu	TCY
26. SINO-TIBETAN						
26.1. Chinese						
Cantonese	cnt	China	23°N	113°E	Chinese (Yue)	YUH
Chaozhou	chz	China	23°40N	116°40E	dialect of Chinese (Min Nan)	dCFR
Fuzhou	fuz	China, Thailand, Malaysia	26°N	119°30E	dialect of Chinese (Min Dong)	dCDO
Hakka	hak	China	25°N	116°E	Chinese (Hakka)	HAK
Hokkien	hok	China	25°N	118°E	dialect of Min Nan Chinese	dCFR
Kunming	kug	China	25°N	102°45E	dialect of Chinese (Mandarin)	dCHN
Mandarin	mnd	China	34°N	110°E	Chinese (Mandarin)	CHN

[8] Treated by *Ethnologue* as a Tibeto-Burman language.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Taiwanese	taq	Taiwan	25°N	121°30E	dialect of Chinese (Min Nan)	dCFR
Wu (Changzhou)	wuc	China	31°40N	119°55E	dialect of Chinese (Wu)	dWUU
Xiamen	xia	China	24°30N	118°10E	dialect of Chinese (Min Nan)	dCFR

26.2. Tibeto-Burman

26.2.1. Bai

Bai	bai	China	26°N	100°E	Bai	PIQ

26.2.2. Baric

Bodo	bod	India	26°50N	92°E	Bodo	BRX
Chang	chg	Myanmar, India	26°20N	94°45E	Naga (Chang)	NBC
Deuri	deu	India	26°N	90°15E	Deori	DER
Dimasa	dms	India	25°30N	93°E	Dimasa	DIS
Garo	gar	India	25°40N	90°30E	Garo	GRT
Jugli	jug	India	27°30N	96°20E	dialect of Naga (Tase)	dNST
Kachari	kac	India	26°30N	91°E	Kachari	QKC
Kokborok	kok	India	24°20N	92°30E	Kok Borok	TRP
Lungchang	lun	India	27°30N	96°25E	dialect of Naga (Tase)	dNST
Nocte	noc	India	27°N	95°30E	Naga (Nocte)	NJB

26.2.3. Bodic

Amdo	amd	China	35°N	100°E	Amdo	ADX
Athpare	ath	Nepal	27°N	87°20E	Athpariya	APH
Balti	blt	Pakistan, India	35°N	76°E	Balti	BFT
Baragaunle	brl	Nepal	28°50N	83°50E	Baragaunle	BON
Belhare	bel	Nepal	26°58N	87°18E	Belhariye	BYW
Byansi	bya	India	30°10N	80°30E	Byangsi	BEE
Camling	cml	Nepal	27°N	86°40E	Camling	RAB
Chantyal	chn	Nepal	28°35N	83°25E	Chantyal	CHX
Chaudangsi	chd	India	30°N	80°15E	Chaudangsi	CDN
Chepang	cpn	Nepal	27°40N	84°45E	Chepang	CDM
Darmiya	drm	India	30°N	79°45E	Darmiya	DRD
Dhimal	dhm	Nepal	26°40N	87°45E	Dhimal	DHI
Dumi	dmi	Nepal	27°15N	86°40E	Dumi	DUS
Gahri	ghr	India	32°30N	77°E	Gahri	BFU
Gurung	gur	Nepal	28°20N	84°20E	2 varieties: Gurung (Eastern)/ Gurung (Western)	GGN/GVR
Hayu	hay	Nepal	27°15N	86°E	Wayu	VAY
Jad	jad	India	30°50N	79°E	dialect of Tibetan	dTIC
Khaling	khg	Nepal, China, India	27°30N	86°40E	Khaling	KLR
Kham	kmh	Nepal	28°30N	82°45E	5 varieties: Kham (Gamale)/. . .	KGJ/. . .
Kinnauri	knn	India	31°30N	78°E	Kinnauri	KFK
Kulung	klg	Nepal	27°30N	87°E	Kulung	KLE
Ladakhi	lad	India	34°N	78°E	Ladakhi	LBJ
Lhomi	lho	Nepal	27°40N	87°25E	Lhomi	LHM
Limbu	lim	Nepal, India, Bhutan	27°10N	87°45E	Limbu	LIF
Magar	mag	Nepal	28°N	83°E	Magar (Eastern)	MGP
Magar (Syangja)	msy	Nepal	28°05N	83°50E	Magar (Western)	MRD
Marchha	mrc	India	30°15N	79°35E	—	—
Nar-Phu	nph	Nepal	28°40N	84°15E	Nar Phu	NPA
Newari (Dolakha)	nwd	Nepal	27°20N	86°E	dialect of Newari	dNEW
Newari (Kathmandu)	new	Nepal	27°40N	85°30E	dialect of Newari	dNEW
Nyamkad	nyk	India	32°N	78°30E	Tukpa	TPQ
Pattani	ptt	India	32°45N	76°45E	Pattani	LAE
Purki	prk	India	34°40N	76°E	Purik	BXR
Sherpa	she	Nepal	27°45N	86°50E	Sherpa	SCR
Sikkimese	skk	India	27°50N	88°30E	Sikkimese	SIP
Spitian	spi	India	32°15N	78°E	dialect of Tibetan	dTIC
Tamang	tam	Nepal	28°N	85°15E	5 varieties: Tamang (Eastern)/. . .	TAJ/ TGE/. . .
Thakali	thk	Nepal	29°N	83°45E	Thakali	THS
Thangmi	thn	Nepal	27°45N	86°E	Thami	THF
Thulung	thu	Nepal	27°25N	86°30E	Thulung	TDH
Tibetan (Modern Literary)	tmo	China	29°50N	91°30E	dialect of Tibetan	dTIC
Tibetan (Shigatse)	tis	China	29°N	89°E	dialect of Tibetan	dTIC
Tibetan (Standard Spoken)	tib	China	30°N	91°E	dialect of Tibetan	dTIC
Tinani	tni	India	32°20N	77°10E	Tinani	LBF
Tod	tod	India	32°25N	77°15E	Stod Bhoti	SBU
Tshangla	tgl	India, Bhutan	27°50N	92°10E	Tshangla	TSJ
Zhang-Zhung	zha	India	31°35N	78°25E	Jangshung	JNA

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
26.2.4. Burmese-Lolo						
Achang	acn	Myanmar, China	25°N	98°30E	Achang	ACN
Akha	akh	Myanmar, China, Thailand	21°50N	99°50E	Akha	AKA
Arakanese (Marma)	akm	Myanmar, Bangladesh	21°30N	92°30E	Arakanese	MHV
Bisu	biu	Thailand	19°45N	100°E	Bisu	BII
Burmese	brm	Myanmar	21°N	96°E	Burmese	BMS
Hani	han	China	23°N	103°E	Hani	HNI
Jino	jin	China	22°N	101°E	2 varieties: Jinuo (Buyuan)/ Jinuo (Youle)	JIY/JIU
Lahu	lah	Myanmar, China, Thailand	20°N	98°10E	Lahu	LAH
Lalo	lal	China	25°N	100°15E	Yi (Western)	YIW
Lisu	lis	China, Myanmar	26°N	98°E	Lisu	LIS
Maru	mru	Myanmar	23°30N	98°30E	Maru	MHX
Nusu	nus	China	26°45N	99°E	Nusu	NUF
Yi	yi	China	24°N	104°E	6 varieties: Yi (Central)/. . .	YIC/. . .
26.2.5. Jinghpo						
Jingpho	jng	Myanmar	25°25N	97°E	Jingpho	CGP
26.2.6. Karen						
Karen (Bwe)	kbw	Myanmar	19°30N	97°E	Karen (Bwe)	BWE
Karen (Pwo)	kpw	Myanmar, China	13°N	99°10E	2 varieties: Karen (Pwo Eastern)/ Karen (Pwo Western)	KJP/PWO
Karen (Sgaw)	ksg	Myanmar	18°N	97°E	Karen (S'gaw)	KSW
Kayah Li (Eastern)	kyl	Myanmar, Thailand	19°N	97°30E	Kayah (Eastern)	EKY
Phlong	phl	Thailand	15°N	99°E	Karen (Pwo Northern)	PWW
26.2.7. Kuki-Chin-Naga						
Angami	agm	India	25°40N	94°30E	2 varieties: Naga (Angami)/ Naga (Chokri)	NJM/NRI
Ao	ao	India	26°35N	94°40E	Naga (Ao)	NJO
Bawm	baw	Bangladesh, India, Myanmar	22°30N	92°15E	Chin (Bawm)	BGR
Chin (Mara)	cmr	India, Myanmar	23°N	93°10E	Chin (Mara)	MRH
Chin (Siyin)	chs	Myanmar	23°50N	94°E	Chin (Siyin)	CSY
Chin (Tiddim)	cti	Myanmar	23°20N	93°40E	Chin (Tedim)	CTD
Hmar	hma	India	24°10N	93°E	Hmar	HMR
Hyow	hyo	Bangladesh	22°05N	92°25E	Chin (Asho)	CSH
Kabui	kbi	India	24°50N	94°E	Naga (Kabui)	NKF
Khumi	khu	Myanmar	21°55N	92°25E	Chin (Khumi)	CKM
Lai	lai	Myanmar	22°40N	93°40E	Chin (Haka)	CNH
Lotha	lot	India	26°30N	94°15E	Naga (Lotha)	NJH
Meithei	mei	India	24°45N	94°E	Meitei	MNR
Mikir	mik	India	26°20N	93°30E	Mikir	MJW
Mizo	miz	India, Bangladesh	23°10N	92°50E	dialect of Lushai	dLSH
Naga (Mao)	nma	India	25°30N	94°20E	Naga (Mao)	NBI
Naga (Tangkhul)	ngt	India	25°N	94°30E	Naga (Tangkhul)	NMF
Naga (Zeme)	nze	India	25°10N	93°30E	Naga (Zeme)	NZM
Paite	pte	India	24°N	93°10E	Chin (Paite)	PCK
Sema	sem	India	26°N	94°30E	Naga (Sumi)	NSM
Tarao	tao	India	24°15N	94°10E	Naga (Tarao)	TRO
Thadou	thd	India, Myanmar	24°25N	93°55E	Chin (Thado)	TCZ
26.2.8. Lepcha						
Lepcha	lep	India, Nepal, Bhutan	27°10N	88°30E	Lepcha	LEP
26.2.9. Mirish						
Apatani	apt	India	27°30N	93°45E	Apatani	APT
Bokar	boq	India	28°30N	94°40E	dialect of Adi	dADI
Bori	boj	India	28°20N	94°45E	dialect of Adi	dADI
Bugun	bgn	India	27°30N	92°35E	—	—
Digaro	dig	India	28°25N	96°E	Digaro	MHU
Gallong	gal	India	28°10N	94°40E	dialect of Adi	dADI
Idu	idu	India	29°20N	95°50E	Idu	CLK
Miju	mij	India	29°N	96°E	Miju	MXJ
Milang	mil	India	28°30N	95°15E	dialect of Adi	dADI
Mising	msg	India	28°30N	94°50E	Miri	MRG
Nishi	nis	India	27°30N	93°30E	Nisi	DAP
Sherdukpen	shd	India	27°N	92°30E	Sherdukpen	SDP
Sulung	slg	India	28°20N	93°15E	Sulung	SUV
26.2.10. Naxi						
Naxi	nax	China	27°30N	100°E	Naxi	NBF

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
26.2.11. Nungish						
Anong	ano	Myanmar, China	27°20N	98°45E	Nung	NUN
Dulong	dul	China	28°N	98°20E	Drung	DUU
Rawang	raw	Myanmar	27°30N	97°30E	Rawang	RAW
26.2.12. Qiangic						
Pumi	pum	China	28°N	101°E	2 varieties: Pumi (Northern)/ Pumi (Southern)	PMI/PUS
Qiang	qia	China	32°N	102°40E	2 varieties: Qiang (Northern)/ Qiang (Southern)	CNG/QMR
26.2.13. rGyalrong						
Jiarong	jia	China	31°30N	102°E	Jiarong	JYA
27. HMONG-MIEN						
Hmong Daw	hmd	China, Vietnam	26°N	105°E	Hmong Daw	MWW
Hmong Njua	hmo	China	28°N	105°E	Hmong Njua	BLU
Mien	mie	China	25°N	111°E	Iu Mien	IUM
Mun	mnu	China	19°N	110°E	Kim Mun	MJI
Punu	pnu	China	24°N	107°E	5 varieties: Bunu (Younuo)/. . .	BUH/. . .
28. TAI-KADAI						
28.1. Kadai						
Gelao	gla	Vietnam	22°55N	105°30E	Gelao	KKF
Lachi	lch	Vietnam	22°40N	104°50E	Lachi	LBT
Laha	lha	Vietnam	21°35N	103°55E	Laha	LHA
Lakkia	lkk	China	24°05N	110°10E	Lakkia	LBC
28.2. Kam-Tai						
Dong	don	China	27°N	109°E	2 varieties: Dong (Northern)/ Dong (Southern)	DOC/KMC
Kam (Zhanglu)	kzh	China	26°N	108°30E	dialect of Dong (Northern)	dDOC
Lao	lao	Laos, Thailand	18°N	103°E	Lao	NOL
Lü	lu	China, Laos, Myanmar	22°N	100°40E	Lü	KHB
Mulao	mul	China	25°N	108°E	Mulao	MLM
Nung (in Vietnam)	nun	Vietnam	21°55N	106°25E	Nung	NUT
Saek	sae	Laos, Thailand	17°25N	104°45E	Saek	SKB
Shan	sha	Myanmar	22°N	98°E	Shan	SJN
Sui	sui	China	26°N	107°30E	Sui	SWI
Thai	tha	Thailand	16°N	101°E	Thai	THJ
Yay	yay	Vietnam	22°25N	104°45E	Nhang	NHA
Zhuang (Northern)	zhn	China	24°N	108°E	Zhuang (Northern)	CCX
29. AUSTRO-ASIATIC						
29.1. Mon-Khmer						
29.1.1. Aslian						
Jahai	jah	Malaysia	5°45N	101°30E	Jehai	JHI
Mah Meri	mhm	Malaysia	2°50N	101°30E	Besisi	MHE
Semai	smj	Malaysia	4°20N	101°40E	Semai	SEA
Semelai	sml	Malaysia	3°N	103°E	Semelai	SZA
Temiar	tmr	Malaysia	5°N	101°30E	Temiar	TMH
29.1.2. Bahnaric						
Bahnar (Plei Bong-Mang Yang)	bpb	Vietnam	13°50N	108°20E	dialect of Bahnar	dBDQ
Brao	bra	Vietnam	14°10N	107°30E	Brao	BRB
Chrau	chr	Vietnam	10°45N	107°30E	Chrau	CHR
Cua	cua	Vietnam	15°15N	108°30E	Cua	CUA
Halang	hln	Vietnam	14°30N	107°30E	Halang	HAL
Hre	hre	Vietnam	14°40N	108°40E	Hre	HRE
Jeh	jeh	Vietnam, Laos	15°10N	107°50E	Jeh	JEH
Loven	lov	Laos	15°N	106°20E	Laven	LBO
Mnong (Eastern)	mge	Vietnam	12°40N	108°30E	Mnong (Eastern)	MNG
Oi	oi	Laos	14°50N	107°10E	Oy	OYB
Rengao	rng	Vietnam	14°35N	107°50E	Rengao	REN

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Sapuan	sap	Laos	15°10N	106°50E	Sapuan	SPU
Sedang	sed	Vietnam	14°50N	108°E	Sedang	SED
Sre	sre	Vietnam	11°30N	108°E	dialect of Koho	dKPM
Stieng	sti	Vietnam, Cambodia	11°50N	106°45E	2 varieties: Stieng (Budeh)/Stienh (Bulo)	STT/STI
29.1.3. Katuic						
Bru (Eastern)	bru	Laos	17°20N	106°E	Bru (Eastern)	BRU
Bru (Western)	brw	Laos	16°45N	104°45E	Bru (Western)	BRV
Katu	ktu	Vietnam, Laos	15°50N	107°35E	2 varieties: Katu (Western)/ Katu (Eastern)	KUF/KTV
Pacoh	pac	Vietnam	16°25N	107°05E	Pacoh	PAC
29.1.4. Khasian						
Khasi	khs	India	25°30N	92°E	Khasi	KHI
29.1.5. Khmer						
Khmer	khm	Cambodia	12°30N	105°E	Khmer (Central)	KMR
29.1.6. Monic						
Mon	mon	Thailand, Myanmar	14°50N	100°30E	Mon	MNW
Nyah Kur (Tha Pong)	nkt	Thailand	15°40N	101°40E	dialect of Nyahkur	dCBN
29.1.7. Nicobarese						
Nancowry	nnc	India	8°03N	93°30E	dialect of Nicobarese (Central)	dNCB
Nicobarese	nic	India	7°N	93°50E	3 varieties: Nicobarese (Central)/. . .	NCB/. . .
Nicobarese (Car)	nca	India	9°N	93°E	Nicobarese (Car)	CAQ
29.1.8. Palaung-Khmuic						
Khmu'	kmu	Laos	21°N	102°E	Khmu	KJG
Ksingmul	ksi	Vietnam, Laos	21°05N	104°E	Puoc	PUO
Mlabri (Minor)	mlm	Thailand	18°30N	101°E	dialect of Mlabri	dMRA
Palaung	plg	Myanmar	22°40N	96°40E	3 varieties: Palaung (Pale)/. . .	PCE/ RBB/. . .
Parauk	puk	Myanmar, China	23°15N	99°30E	Parauk	PRK
29.1.9. Viet-Muong						
Muong	muo	Vietnam	20°30N	105°05E	Muong	MTQ
Vietnamese	vie	Vietnam	10°30N	106°30E	Vietnamese	VIE
29.2. Munda						
Bhumij	bhu	India	21°20N	86°30E	dialect of Mundari	dMUW
Gadaba	gdb	India	19°N	83°40E	3 varieties: Gadaba (Bodo)/. . .	GBJ/. . .
Ho	ho	India	22°N	86°E	Ho	HOC
Juang	jun	India	21°20N	86°E	Juang	JUN
Kharia	khr	India	22°30N	84°20E	Kharia	KHR
Korku	kku	India	22°30N	78°30E	Korku	KFQ
Mundari	mun	India	23°N	84°40E	Mundari	MUW
Remo	rem	India	18°N	82°E	Bondo	BFW
Santali	stl	India	24°30N	87°E	Santali	SNT
Sora	sor	India	20°N	84°20E	Sora	SRB
30. ANDAMANESE						
30.1. Great Andamanese						
Aka-Biada	akb	India	11°40N	92°30E	dialect of Aka-Bea	dACE
Aka-Cari	akc	India	13°30N	93°E	Aka-Cari	ACI
Aka-Kede	akk	India	12°45N	92°45E	Aka-Kede	AKX
A-Pucikwar	apk	India	12°10N	92°50E	A-Pucikwar	APQ
Great Andamanese	gan	India	12°N	92°40E	10 varieties: A-Pucikwar/. . .	APQ/. . .
30.2. South Andamanese						
Jarawa (in Andamans)	jrw	India	12°N	92°35E	Jarawa	ANQ
Onge	ong	India	10°45N	92°25E	Önge	OON
31. AUSTRONESIAN						
31.1. Atayalic						
Atayal	ata	Taiwan	24°30N	121°20E	Atayal	TAY
Seediq	see	Taiwan	24°10N	121°25E	Taroko	TRV

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
31.2. Central Malayo-Polynesian						
Alune	aln	Indonesia	3°S	128°20E	Alune	ALP
Bima	bim	Indonesia	8°30S	118°30E	Bima	BHP
Buru	buu	Indonesia	3°30S	126°30E	Buru	MHS
Dobel	dob	Indonesia	6°15S	134°40E	Dobel	KVO
Fordata	frd	Indonesia	6°45S	131°30E	Fordata	FRD
Kambera	kam	Indonesia	9°50S	120°10E	Kambera	SMI
Kedang	ked	Indonesia	8°15S	123°45E	Kedang	KSX
Kei	kei	Indonesia	5°50S	132°55E	Kei	KEI
Ke'o	keo	Indonesia	8°50S	121°15E	Ke'o	XXK
Kisar	ksr	Indonesia	8°05S	127°07E	Kisar	KJE
Kola	koo	Indonesia	5°30S	134°30E	Kola	KVV
Lamaholot	lmh	Indonesia	8°15S	122°55E	Lamaholot	SLP
Larike	lrk	Indonesia	3°45S	127°55E	Larike-Wakasihu	ALO
Leti	let	Indonesia	8°12S	127°40E	Leti	LTI
Manggarai	mng	Indonesia	8°30S	120°20E	Manggarai	MQY
Ngad'a	ngd	Indonesia	8°50S	121°E	Ngad'a	NXG
Nuaulu	nua	Indonesia	3°15S	129°10E	2 varieties: Nuaulu (North)/ Nuaulu (South)	NNI/NXL
Paulohi	plh	Indonesia	3°15S	128°45E	Paulohi	PLH
Roti	rti	Indonesia	10°40S	123°15E	Rote	ROT
Sawu	saw	Indonesia	11°30S	121°55E	Sabu	HVN
Selaru	sru	Indonesia	8°12S	131°E	Selaru	SLU
Sika	sik	Indonesia	8°42S	122°15E	Sika	SKI
Tarangan (West)	tgw	Indonesia	6°30S	134°10E	Tarangan (West)	TXN
Tetun	ttn	East Timor	9°S	126°E	Tetun	TTM
Timorese	tse	Indonesia	9°40S	124°30E	Atoni	TMR
Tugun	tgn	Indonesia	7°43S	126°45E	Tugun	TZN
Wetan	wet	Indonesia	8°10S	128°E	Luang	LEX
Yamdena	ymd	Indonesia	7°30S	131°30E	Yamdena	JMD
31.3. Eastern Malayo-Polynesian						
31.3.1. Oceanic						
Adzera	adz	Papua New Guinea	6°15S	146°15E	Adzera	AZR
Ajië	aji	New Caledonia	21°20S	165°30E	Ajië	AJI
Ala'ala	all	Papua New Guinea	8°55S	146°45E	Nara	NRZ
Amara	ama	Papua New Guinea	5°40S	148°30E	Amara	AIE
Ambae (Lolovoli Northeast)	aml	Vanuatu	15°25S	167°53E	Ambae (East)	OMB
Anejom	anj	Vanuatu	20°12S	169°48E	Aneityum	ATY
Araki	ark	Vanuatu	15°40S	166°55E	Araki	AKR
'Are'are	arx	Solomon Islands	9°15S	161°10E	'Are'are	ALU
Aribwatsa	ari	Papua New Guinea	6°43S	147°E	Aribwatsa	LAZ
Arop-Lokep	alk	Papua New Guinea	5°23S	147°03E	Arop-Lokep	APR
Arosi	aro	Solomon Islands	10°15S	161°20E	Arosi	AIA
Atchin	atc	Vanuatu	16°S	167°20E	Uripiv-Wala-Rano-Atchin	UPV
Aulua	aul	Vanuatu	16°20S	167°40E	Aulua	AUL
Bali-Vitu	bvi	Papua New Guinea	4°54S	149°07E	2 varieties: Muduapa/Uneapa	WIV/BBN
Banoni	bnn	Papua New Guinea	6°25S	155°15E	Banoni	BCM
Buma	bum	Solomon Islands	11°38S	166°50E	Teanu	TKW
Carolinian	crl	Northern Mariana Islands	15°12N	145°45E	Carolinian	CAL
Cèmuhî	cem	New Caledonia	20°50S	165°10E	Cemuhî	CAM
Cheke Holo	ckh	Solomon Islands	8°20S	159°40E	Cheke Holo	MRN
Chuukese	cuu	Micronesia	7°20N	151°45E	Chuukese	TRU
Drehu	dre	New Caledonia	21°S	167°15E	Dehu	DEU
Erromangan	err	Vanuatu	18°50S	169°10E	Sie	ERG
Fijian	fij	Fiji	17°50S	178°E	Fijian	FJI
Fijian (Wayan)	fiw	Fiji	17°S	177°18E	dialect of Fijian (Western)	dWYY
Futuna (East)	fue	Wallis and Futuna	14°20S	178°10W	Futuna (East)	FUD
Futuna-Aniwa	fut	Vanuatu	19°32S	170°13E	Futuna-Aniwa	FUT
Gapapaiwa	gap	Papua New Guinea	9°45S	149°50E	Gapapaiwa	PWG
Gela	gel	Solomon Islands	9°05S	160°15E	Gela	NLG
Gumawana	gmw	Papua New Guinea	9°15S	150°50E	Gumawana	GVS
Halia	hal	Papua New Guinea	5°15S	154°40E	Halia	HLA
Hawaiian	haw	United States	19°35N	155°30W	Hawaiian	HWI
Hoava	hoa	Solomon Islands	8°05S	157°30E	Hoava	HOA
Iaai	iaa	New Caledonia	20°25S	166°35E	Iaai	IAI
Iduna	idn	Papua New Guinea	9°20S	150°15E	Iduna	VIV
Jabêm	jab	Papua New Guinea	6°35S	147°47E	Yabem	JAE
Kairiru	krr	Papua New Guinea	3°20S	143°35E	Kairiru	KXA
Kaliai-Kove	kkv	Papua New Guinea	5°35S	149°40E	2 varieties: Lusi/Kove	KHL/KVC
Kapingamarangi	kpn	Micronesia	1°03N	154°45E	Kapingamarangi	KPG

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Kara (in Papua New Guinea)	kra	Papua New Guinea	2°50S	151°07E	Kara	LEU
Kaulong	kau	Papua New Guinea	6°10S	149°40E	Kaulong	PSS
Kela (Apoze)	kap	Papua New Guinea	7°25S	147°05E	dialect of Kela	dKCL
Kele	kel	Papua New Guinea	2°05S	147°05E	Kele	SBC
Kilivila	klv	Papua New Guinea	8°30S	151°05E	Kilivila	KIJ
Kiribati	krb	Kiribati	1°20N	173°E	Kiribati	GLB
Kokota	kkt	Solomon Islands	8°S	159°08E	Kokota	KKK
Kosraean	kos	Micronesia	5°18N	163°E	Kosraean	KSI
Kwaio	kwa	Solomon Islands	8°57S	161°E	Kwaio	KWD
Kwamera	kwr	Vanuatu	19°35S	169°25E	Kwamera	TNK
Lau	lau	Solomon Islands	9°35S	161°30E	Lau	LLU
Lenakel	len	Vanuatu	19°27S	169°15E	Lenakel	TNL
Lewo	lew	Vanuatu	16°45S	168°20E	Lewo	LWW
Longgu	lgu	Solomon Islands	9°45S	160°40E	Longgu	LGU
Loniu	lon	Papua New Guinea	2°04S	147°20E	Loniu	LOS
Lonwolwol	lnw	Vanuatu	16°13S	168°E	Lonwolwol	CRC
Lou	lou	Papua New Guinea	2°23S	147°19E	Lou	LOJ
Luangiua	lga	Solomon Islands	5°20S	159°25E	Ongtong Java	LUN
Mae	mae	Vanuatu	16°S	167°15E	Mae	MME
Maisin	msn	Papua New Guinea	9°30S	149°10E	Maisin	MBQ
Maleu	mlu	Papua New Guinea	5°43S	148°25E	Maleu-Kilenge	MGL
Manam	mnm	Papua New Guinea	4°S	145°E	Manam	MVA
Mangap-Mbula	mmb	Papua New Guinea	5°40S	148°05E	Mbula	MNA
Maori	mao	New Zealand	40°S	176°E	Maori	MBF
Marquesan	mrq	French Polynesia	8°55S	140°05W	2 varieties: Marquesan (North)/ Marquesan (South)	MRQ/QMS
Marshallese	msh	Marshall Islands	7°07N	171°03E	Marshallese	MZM
Matukar	mtk	Papua New Guinea	4°55S	145°45E	Matukar	MJK
Mekeo	mke	Papua New Guinea	8°20S	146°30E	Mekeo	MEK
Minaveha	mnv	Papua New Guinea	9°36S	150°28E	Minaveha	MVN
Mokilese	mok	Micronesia	6°40N	159°45E	Mokilese	MNO
Mono-Alu	moa	Solomon Islands	7°03S	155°45E	Mono	MTE
Motu	mtu	Papua New Guinea	9°20S	147°E	Motu	MEU
Musom	msm	Papua New Guinea	6°35S	147°E	Musom	MSU
Mussau	mus	Papua New Guinea	1°27S	149°37E	Mussau-Emira	EMI
Muyuw	muy	Papua New Guinea	9°05S	152°45E	Muyuw	MYW
Mwotlap	mwo	Vanuatu	13°35S	167°35E	Motlav	MLV
Nakanai	nak	Papua New Guinea	5°35S	150°35E	Nakanai	NAK
Nalik	nal	Papua New Guinea	2°57S	151°20E	Nalik	NAL
Nambas (Big)	nbb	Vanuatu	16°05S	167°12E	Nambas (Big)	NMB
Nanumea	nnm	Tuvalu	5°40S	176°07E	dialect of Tuvaluan	dELL
Nauruan	nau	Nauru	0°30S	166°55E	Nauruan	NRU
Nehan	neh	Papua New Guinea	4°30S	154°12E	Nehan	NSN
Nelemwa	nel	New Caledonia	20°15S	164°05E	Kumak	NEE
Nengone	nne	New Caledonia	21°30S	168°E	Nengone	NEN
Nguna	ngu	Vanuatu	16°55S	168°30E	dialect of Efate (North)	dLLP
Niuafo'ou	nif	Tonga	15°34S	175°37W	Niuafo'ou	NUM
Niuean	niu	New Zealand	19°03S	170°07W	Niue	NIQ
Nukuoro	nkr	Micronesia	3°50N	154°55E	Nukuoro	NKR
Nyelâyu	nyl	New Caledonia	20°20S	164°20E	Nyâlayu	YLY
Paamese	pms	Vanuatu	16°30S	168°15E	Paama	PMA
Paita	pta	New Caledonia	22°05S	166°30E	Dumbea	DUF
Patep	pat	Papua New Guinea	6°55S	146°35E	Patep	PTP
Patpatar	ptp	Papua New Guinea	3°46S	152°30E	Patpatar	GFK
Pileni	pil	Solomon Islands	10°13S	166°13E	Pileni	PIV
Pingilapese	pgl	Micronesia	6°13N	160°42E	Pingelapese	PIF
Po-Ai	poa	New Caledonia	20°40S	164°50E	Fwâi	FWA
Pohnpeian	poh	Micronesia	6°53N	158°15E	Pohnopeian	PNF
Puluwat	pul	Micronesia	7°20N	149°20E	Puluwatese	PUW
Raga	rag	Vanuatu	15°30S	168°10E	Hano	LML
Rapanui	rap	Chile	27°S	109°W	Rapa Nui	PBA
Rennellese	rnn	Solomon Islands	11°37S	160°15E	Rennell	MNV
Roro	ror	Papua New Guinea	8°45S	146°35E	Roro	RRO
Rotuman	rot	Fiji	12°30S	177°04E	Rotuman	RTM
Roviana	rov	Solomon Islands	8°15S	157°20E	Roviana	RUG
Sa'a	saa	Solomon Islands	9°37S	161°27E	Sa'a	APB
Sakao	sak	Vanuatu	15°10S	167°05E	Sakao	SKU
Saliba (in Papua New Guinea)	slb	Papua New Guinea	10°35S	150°43E	Saliba	SBE
Samoan	sam	Samoa, American Samoa	13°55S	171°50W	Samoan	SMY
Siar	sir	Papua New Guinea	4°40S	152°55E	Lak	SJR
Sinaugoro	sng	Papua New Guinea	9°50S	147°50E	Sinaugoro	SNC
Sio	sio	Papua New Guinea	5°57S	147°20E	Sio	SIO

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Sisiqa	sis	Solomon Islands	7°S	156°50E	dialect of Babatana	dBBQ
Sobei	sob	Indonesia	1°55S	138°45E	Sobei	SOB
Sonsorol-Tobi	son	Palau	5°20N	132°15E	Sonsorol	SOV
Southeast Ambrym	sea	Vanuatu	16°18S	168°13E	Southeast Ambrym	TVK
Sudest	sud	Papua New Guinea	11°30S	153°27E	Sudest	TGO
Sursurunga	sur	Papua New Guinea	4°S	152°46E	Sursurunga	SGZ
Tahitian	tah	French Polynesia	17°40S	149°35W	Tahitian	THT
Taiof	taf	Papua New Guinea	5°32S	154°38E	Saposa	SPS
Takia	tak	Papua New Guinea	4°40S	146°E	Takia	TBC
Tamabo	tmm	Vanuatu	15°40S	167°10E	dialect of Malo	dMLA
Tangga	tga	Papua New Guinea	3°28S	153°12E	Tangga	TGG
Tanna (Southwest)	tns	Vanuatu	19°34S	169°20E	Tanna (Southwest)	NWI
Tawala	taw	Papua New Guinea	10°20S	150°40E	Tawala	TBO
Teop	teo	Papua New Guinea	5°40S	155°E	Teop	TIO
Tigak	tgk	Papua New Guinea	2°43S	150°48E	Tigak	TGC
Tinrin	tin	New Caledonia	21°40S	165°45E	Tiri	CIR
Tobati	tbt	Indonesia	2°35S	140°40E	Tobati	TTI
Tokelauan	tke	Tokelau	9°S	172°W	Tokelauan	TOK
Tolai	tla	Papua New Guinea	4°15S	152°05E	Kuanua	KSD
Tongan	tng	Tonga	21°10S	175°10W	Tongan	TOV
Toqabaqita	toq	Solomon Islands	8°25S	160°35E	To'abaita	MLU
Tuamotuan	tua	French Polynesia	17°S	144°W	Tuamotuan	PMT
Tumleo	tum	Papua New Guinea	3°05S	142°25E	Tumleo	TMQ
Tungak	tnk	Papua New Guinea	2°33S	150°15E	Tungag	LCM
Tuvaluan	tvl	Tuvalu	8°30S	179°10E	Tuvaluan	ELL
Ulithian	uli	Micronesia	9°55N	139°35E	Ulithian	ULI
Ura	ura	Vanuatu	18°40S	169°05E	Ura	UUR
Wallisian	wll	Wallis and Futuna	13°18S	176°12W	Wallisian	WAL
Wedau	wed	Papua New Guinea	10°10S	150°10E	Wedau	WED
Woleaian	wol	Micronesia	7°20N	143°50E	Woleaian	WOE
Xârâcùù	xar	New Caledonia	21°40S	166°E	Xârâcùù	ANE
31.3.2. South Halmahera—West New Guinea						
Ambai	amq	Indonesia	1°50S	136°30E	Ambai	AMK
Biak	bik	Indonesia	1°S	136°E	Biak	BHW
Buli (in Indonesia)	bui	Indonesia	1°N	128°30E	Buli	BZQ
Irarutu	irr	Indonesia	3°S	133°30E	Irarutu	IRH
Ma'ya	mya	Indonesia	1°15S	130°55E	Ma'ya	SLZ
Mor	mor	Indonesia	3°S	135°45E	Mor	MHZ
Sawai	swi	Indonesia	0°30N	128°E	Sawai	SZW
Taba	tab	Indonesia	0°N	127°30E	Makian (East)	MKY
Waropen	wrp	Indonesia	2°20S	136°35E	Waropen	WRP
31.4. Paiwanic						
Amis	ami	Taiwan	23°20N	121°20E	Amis	ALV
Paiwan	pai	Taiwan	22°30N	120°50E	Paiwan	PWN
Siraiya	sry	Taiwan	23°N	120°15E	Siraiya	FOS
Thao	thw	Taiwan	23°53N	120°55E	Thao	SSF
31.5. Tsouic						
Rukai	ruk	Taiwan	22°50N	120°50E	Rukai	DRU
Tsou	tso	Taiwan	23°30N	120°45E	Tsou	TSY
31.6. Western Malayo-Polynesian						
31.6.1. Borneo						
Begak-Ida'an	beg	Malaysia	5°15N	118°50E	dialect of Ida'an	dDBJ
Biatah	bit	Malaysia, Indonesia	1°15N	110°E	Biatah	BTH
Kadazan	kdz	Malaysia	5°55N	117°10E	3 varieties: Kadazan (Coastal)/. . .	KZJ/. . .
Kayan (Baram)	kbr	Malaysia	3°30N	114°30E	Kayan (Baram)	KYS
Kelabit	klt	Malaysia	3°40N	115°25E	Kelabit	KZI
Kenyah (Uma' Lung)	keu	Indonesia	3°N	115°50E	—	—
Lun Dayeh	lud	Indonesia	4°N	115°55E	Lundayeh	LND
Ma'anyan	myn	Indonesia	2°S	115°E	Ma'anyan	MHY
Malagasy	mal	Madagascar	20°S	47°E	Malagasy	MEX
Melanau	mel	Malaysia	2°30N	111°30E	Melanau	MEL
Mentuh Tapuh	mta	Malaysia	1°15N	111°E	dialect of Bukar Sadong	dSDO
Narom	nrm	Malaysia	4°25N	114°E	Narom	NRM
Oloh Mangtangai	olm	Indonesia	0°50S	113°E	dialect of Ngaju	dNIJ
Pulopetak	plp	Indonesia	1°25N	109°50E	dialect of Ngaju	dNIJ
Sa'ban	sab	Indonesia	3°40N	115°40E	Sa'ban	SNV

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Semandang	smd	Indonesia	0°45S	110°30E	Semandang	SDN
Tatana'	tat	Malaysia	5°30N	115°30E	Tatana	TXX
Timugon	tim	Malaysia	5°N	116°E	Timugon Murut	TIH
31.6.2. Chamorro						
Chamorro	cha	Guam	13°27N	144°45E	Chamorro	CJD
31.6.3. Meso-Philippine						
Aklanon	akl	Philippines	11°35N	122°20E	Aklanon	AKL
Batak	bat	Philippines	10°15N	119°05E	Batak	BTK
Bikol	bkl	Philippines	13°20N	123°30E	8 varieties: Bicolano (Central)/. . .	BKL/. . .
Cebuano	ceb	Philippines	10°N	124°E	Cebuano	CEB
Hanunóo	hnn	Philippines	12°20N	121°15E	Hanunoo	HNN
Hiligaynon	hil	Philippines	10°15N	123°E	Hiligaynon	HIL
Mamanwa	mmn	Philippines	9°25N	125°30E	Mamanwa	MMN
Tagalog	tag	Philippines	15°N	121°E	Tagalog	TGL
Tagbanwa (Aborlan)	tgb	Philippines	9°30N	118°30E	Tagbanwa (Aborlan)	TBW
Tausug	tsg	Philippines, Indonesia, Malaysia	6°N	121°E	Tausug	TSG
Waray-Waray	wwy	Philippines	12°N	125°E	Waray-Waray	WRY
31.6.4. Northern Philippines						
Agta (Central)	agc	Philippines	17°58N	121°50E	Agta (Central Cagayan)	AGT
Balangao	blg	Philippines	17°10N	121°10E	Balangao	BLW
Bontok	btk	Philippines	17°05N	120°55E	Bontok (Central)	BNC
Dumagat (Casiguran)	dca	Philippines	16°20N	122°E	Agta (Casiguran Dumagat)	DGC
Ibanag	ibn	Philippines	17°30N	121°40E	Ibanag	IBG
Ifugao (Batad)	ifu	Philippines	16°50N	121°05E	Ifugao (Batad)	IFB
Ilocano	ilo	Philippines	16°N	121°E	Ilocano	ILO
Isnag	isn	Philippines	18°15N	121°E	Isnag	ISD
Itawis	itw	Philippines	17°45N	121°30E	Itawit	ITV
Ivatan	iva	Philippines	20°30N	122°E	Ivatan	IVV
Ivatan (Southern)	ivs	Philippines	20°20N	121°50E	dialect of Ivatan	dIVV
Kankanay	kky	Philippines	16°45N	120°35E	Kankanaey	KNE
Kapampangan	kpm	Philippines	15°20N	120°30E	Pampangan	PMP
Karao	kao	Philippines	16°30N	120°50E	Karao	KYJ
Pangasinan	pnn	Philippines	15°55N	120°20E	Pangasinan	PNG
Yami	ymi	Taiwan	22°N	121°30E	Yami	YMI
31.6.5. Palauan						
Palauan	pal	Palau	7°30N	134°35E	Palauan	PLU
31.6.6. Sama-Bajaw						
Bajau	baj	Indonesia	4°20S	123°E	Bajau (Indonesian)	BDL
Bajau (Semporna)	bjs	Malaysia, Philippines	4°30N	118°30E	dialect of Sama (Southern)	dSIT
Sama (Balangingi)	sbg	Philippines	6°10N	121°50E	Sama (Balangingi)	SSE
Yakan	ykn	Philippines	6°30N	122°E	Yakan	YKA
31.6.7. South Mindanao						
Bilaan	biq	Philippines	6°N	125°20E	2 varieties: Blaan (Sarangani)/ Blaan (Koronadal)	BIS/SMK
Tboli	tbo	Philippines	6°10N	124°30E	Tboli	TBL
Tiruray	try	Philippines	6°45N	124°10E	Tiruray	TIY
31.6.8. Southern Philippines						
Binukid	bkd	Philippines	8°15N	124°50E	Binukid	BKD
Kulamanen	kuz	Philippines	7°45N	125°E	dialect of Manobo (Matigsalug)	dMBT
Magindanao	mgn	Philippines	6°50N	124°30E	Magindanaon	MDH
Manobo (Western Bukidnon)	mwb	Philippines	7°40N	124°45E	Manobo (Western Bukidnon)	MBB
Maranao	mrn	Philippines	7°50N	124°15E	Maranao	MRW
31.6.9. Sulawesi						
Aralle-Tabulahan	atb	Indonesia	2°45S	119°15E	Aralle-Tabulahan	ATQ
Balantak	blk	Indonesia	0°50S	123°15E	Balantak	BLZ
Bambam	bbm	Indonesia	3°05S	119°05E	Bambam	PTU
Banggai	bgg	Indonesia	1°25S	123°10E	Banggai	BGZ
Bantik	bnt	Indonesia	1°25N	124°45E	Bantik	BNQ
Bugis	bug	Indonesia	4°S	120°E	Bugis	BPR
Da'a	daa	Indonesia	1°S	119°35E	dialect of Kaili (Da'a)	dKZF
Gorontalo	grt	Indonesia	0°30N	122°E	Gorontalo	GRL
Kaili	kli	Indonesia	1°10S	120°E	Kaili (Ledo)	LEW

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Konjo	kjo	Indonesia	5°28S	120°20E	Konjo (Coastal)	KJC
Lauje	lje	Indonesia	0°30S	120°E	Lauje	LAW
Makassar	mks	Indonesia	5°25S	119°35E	Makasar	MSR
Mamasa	mms	Indonesia	3°05S	119°40E	Mamasa	MQJ
Mandar	mnr	Indonesia	3°25S	119°05E	Mandar	MHN
Mongondow	mgd	Indonesia	0°40N	124°E	Mongondow	MOG
Muna	mna	Indonesia	5°S	122°30E	Muna	MYN
Napu	npu	Indonesia	1°20S	120°25E	Napu	NAP
Padoe	pad	Indonesia	2°20S	121°20E	Padoe	PDO
Pamona	pna	Indonesia	1°45S	120°50E	Pamona	BCX
Sangir	sgr	Indonesia	3°30N	125°30E	Sangir	SAN
Selayar	sly	Indonesia	6°S	120°30E	Selayar	SLY
Talaud	tld	Indonesia	4°15N	126°45E	Talaud	TLD
Tondano	tno	Indonesia	1°15N	125°E	Tondano	TDN
Tontemboan	tnt	Indonesia	1°05N	124°30E	Tontemboan	TNT
Toraja	trd	Indonesia	3°S	119°45E	Toraja-Sa'dan	SDA
Toratán	tor	Indonesia	1°05N	124°50E	Ratahan	RTH
Tukang Besi	tuk	Indonesia	5°30S	123°30E	2 varieties: Tukangbesi (North)/ Tukangbesi (South)	KHC/BHQ
Uma	uma	Indonesia	1°50S	120°E	Uma	PPK
Wolio	wlo	Indonesia	5°30S	122°45E	Wolio	WLO
31.6.10. Sundic						
Acehnese	ace	Indonesia	5°30N	95°30E	Aceh	ATJ
Alas	alx	Indonesia	3°40N	97°50E	Batak Alas-Kluet	BTZ
Balinese	bal	Indonesia	8°20S	115°15E	Bali	BZC
Batak (Karo)	bkr	Indonesia	3°15N	98°15E	Batak (Karo)	BTX
Batak (Toba)	bto	Indonesia	2°30N	99°E	Batak (Toba)	BBC
Cham (Eastern)	cme	Vietnam	11°20N	108°30E	Cham (Eastern)	CJM
Cham (Western)	chw	Cambodia	12°N	105°30E	Cham (Western)	CJA
Chru	crh	Vietnam	11°30N	108°30E	Chru	CJE
Embaloh	eml	Indonesia	1°10N	112°20E	Embaloh	EMB
Enggano	eno	Indonesia	5°25S	102°15E	Enggano	ENO
Gayo	gay	Indonesia	4°30N	97°30E	Gayo	GYO
Iban	iba	Malaysia	2°N	112°E	Iban	IBA
Indonesian	ind	Indonesia	0°S	106°E	Indonesian	INZ
Indonesian (Irianese)	iir	Indonesia	2°30S	140°40E	dialect of Indonesian	dINZ
Indonesian (Jakarta)	inj	Indonesia	6°10S	106°45E	dialect of Indonesian	dINZ
Javanese	jav	Indonesia	7°S	111°E	Javanese	JAN
Kayu Agung	kag	Indonesia	3°10S	104°55E	Kayu Agung	VKY
Ketapang	ktp	Indonesia	1°52S	110°E	dialect of Malay	dMLI
Kualan	kua	Indonesia	0°25S	110°25E	—	—
Kutai	kuy	Indonesia	0°20S	116°40E	Malay (Tenggarong Kutai)	VKT
Lampung	lmp	Indonesia	5°S	105°E	Lampung	LJP
Madurese	mdr	Indonesia	7°S	113°30E	Madura	MHJ
Malay	mly	Malaysia	3°N	102°E	Malay	MLI
Malay (Kuala Lumpur)	myk	Malaysia	3°10N	101°42E	dialect of Malay	dMLI
Malay (Ulu Muar)	mmu	Malaysia	2°40N	102°30E	Negeri Sembilan Malay	ZMI
Manadonese	mds	Indonesia	1°30N	124°50E	Malay (Manadonese)	XMM
Mentawai	mnt	Indonesia	1°30S	99°E	Mentawai	MWV
Minangkabau	min	Indonesia	1°S	101°E	Minangkabau	MPU
Nias	nia	Indonesia	1°N	97°45E	Nias	NIP
Rade	rad	Vietnam	13°N	108°30E	Rade	RAD
Rejang	rej	Indonesia	3°25S	102°30E	Rejang	REJ
Roglai (Northern)	rgn	Vietnam	12°10N	108°50E	Roglai (Northern)	ROG
Sikule	skl	Indonesia	2°40N	96°E	Sikule	SKH
Simeulue	sim	Indonesia	2°30N	96°15E	Simeulue	SMR
Sundanese	sun	Indonesia	7°S	107°E	Sunda	SUO
Tsat	tst	China	18°15N	109°30E	Tsat	HUQ
Urak Lawoi'	url	Thailand	8°N	98°20E	Urak Lawoi'	URK
31.6.11. Yapese						
Yapese	yap	Micronesia	9°35N	138°10E	Yapese	YPS
32. WEST PAPUAN						
32.1. Hatam						
Hatam	hat	Indonesia	1°S	134°E	Hatam	HAD
32.2. Kebar						
Mpur	mpu	Indonesia	0°45S	133°15E	Mpur	AKC

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
32.3. North Halmaheran						
Galela	gll	Indonesia	2°25N	128°20E	Galela	GBI
Pagu	pag	Indonesia	1°15N	127°45E	Pagu	PGU
Sahu	sah	Indonesia	1°10N	127°30E	Sahu	SUX
Tabaru	tbr	Indonesia	1°30N	127°15E	Tabaru	TBY
Ternate	trt	Indonesia	0°50N	127°15E	Ternate	TFT
Tidore	tid	Indonesia	0°45N	127°30E	Tidore	TVO
Tobelo	tlo	Indonesia	1°30N	128°30E	Tobelo	TLB
West Makian	wma	Indonesia	0°30N	127°35E	West Makian	MQS
32.4. North-Central Bird's Head						
Abun	abu	Indonesia	0°30S	132°30E	Abun	KGR
Maybrat	may	Indonesia	1°20S	132°30E	Mai Brat	AYZ
32.5. West Bird's Head						
Tehit	tht	Indonesia	1°30S	132°E	Tehit	KPS
33. EAST BIRD'S HEAD						
Meyah	mea	Indonesia	1°10S	133°30E	Meyah	MEJ
Sougb	sgb	Indonesia	1°30S	134°E	Manikion	MNX
34. WEST BOMBERAI[9]						
Iha	iha	Indonesia	2°55S	132°15E	Iha	IHP
35. YAWA[10]						
Yawa	yaw	Indonesia	1°45S	136°15E	Yawa	YVA
36. LOWER MAMBERAMO						
Warembori	wrm	Indonesia	1°40S	137°35E	Warembori	WSA
37. EAST GEELVINK BAY[11]						
Bauzi	bzi	Indonesia	2°30S	137°30E	Bauzi	PAU
38. LAKES PLAIN[12]						
Doutai	dou	Indonesia	3°20S	138°10E	Doutai	TDS
Iau	iau	Indonesia	3°10S	137°30E	Iau	TMU
Kirikiri	kkr	Indonesia	3°S	137°10E	Kirikiri	KIY
Obokuitai	obk	Indonesia	3°S	138°E	Obokuitai	AFZ
Sikaritai	skr	Indonesia	2°55S	138°15E	Sikaritai	TTY
39. KWERBA[13]						
Isirawa	isi	Indonesia	1°50S	138°30E	Isirawa	SRL
Kwerba	kwb	Indonesia	2°15S	138°30E	Kwerba	KWE
40. TOR[14]						
Berik	brk	Indonesia	2°15S	138°50E	Berik	BER
Orya	ory	Indonesia	2°45S	139°50E	Orya	URY

[9] Treated by *Ethnologue* as a branch of Trans-New Guinea.

[10] Treated by *Ethnologue* as belonging to a family Geelvink Bay, along with Lakes Plain and East Geelvink Bay languages. These three groups are treated here as three separate families.

[11] Treated by *Ethnologue* as belonging to a family Geelvink Bay, along with Lakes Plain and Yawa. These three groups are treated here as three separate families.

[12] Treated by *Ethnologue* as belonging to a family Geelvink Bay, along with East Geelvink Bay and Yawa. These three groups are treated here as three separate families.

[13] Treated by *Ethnologue* as a subbranch of a Dani-Kwerba group within Trans-New Guinea. While the Dani languages are treated here as Trans-New Guinea, the Kwerba languages are not.

[14] Treated by *Ethnologue* as a branch of Trans-New Guinea.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
41. NIMBORAN[15]						
Kemtuik	kmt	Indonesia	2°40S	140°20E	Kemtuik	KMT
Nimboran	nim	Indonesia	2°30S	140°10E	Nimboran	NIR
42. SENTANI[16]						
Sentani	snt	Indonesia	2°35S	140°35E	Sentani	SET
Tabla	tbl	Indonesia	2°27S	140°25E	Tabla	TNM
43. MORWAP[17]						
Elseng	els	Indonesia	3°S	140°30E	Elseng	MRF
44. KAURE[18]						
Kaure	kaj	Indonesia	3°25S	140°05E	Kaure	BPP
45. SKO[19]						
45.1. Eastern Sko[20]						
Poko-Rawo	pok	Papua New Guinea	2°50S	141°35E	Rawo	RWA
Womo	wom	Papua New Guinea	2°55S	141°50E	—	—
45.2. Krisa[21]						
I'saka	isa	Papua New Guinea	2°50S	141°17E	Krisa	KRO
45.3. Western Sko[22]						
Dumo	dum	Papua New Guinea	2°40S	141°18E	Vanimo	VAM
Sko	sko	Indonesia	2°40S	140°55E	Skou	SKV
46. BORDER[23]						
Amanab	amn	Papua New Guinea	3°35S	141°15E	Amanab	AMN
Awyi	awy	Indonesia	2°55S	140°35E	Awyi	AUW
Imonda	imo	Papua New Guinea	3°20S	141°10E	Imonda	IMN
Manem	mem	Indonesia, Papua New Guinea	3°S	140°50E	Manem	JET
Waris	wrs	Papua New Guinea, Indonesia	3°10S	141°E	Waris	WRS
47. SENAGI[24]						
Anggor	ang	Papua New Guinea	3°45S	141°10E	Angor	AGG
Dera	der	Papua New Guinea, Indonesia	3°35S	141°E	Kamberataro	KBV
48. KARKAR-YURI						
Karkar-Yuri	kyr	Papua New Guinea	3°45S	141°05E	Karkar-Yuri	YUJ
49. YALE						
Nagatman	nag	Papua New Guinea	3°45S	141°30E	Yale	NCE
50. OKSAPMIN[25]						
Oksapmin	oks	Papua New Guinea	5°10S	142°10E	Oksapmin	OPM

[15] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[16] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[17] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[18] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[19] Foley (2000) uses the expression "Sko Family" to refer specifically to what is treated here as Western Sko.
[20] These languages are treated as belonging to the Krisa group by *Ethnologue*. There is no evidence, however, that these languages are any more closely related to the language I'saka (formerly known as Krisa) than they are to Western Sko languages, and are if anything closer to Western Sko.
[21] The group called Krisa in *Ethnologue* also includes what is called here Eastern Sko.
[22] Called *Vanimo* in *Ethnologue*.
[23] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[24] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[25] Treated by *Ethnologue* as a branch of Trans-New Guinea.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
51. TORRICELLI[26]						
51.1. Kombio-Arapesh						
Arapesh	arp	Papua New Guinea	3°28S	143°10E	2 varieties: Bukiyip/ Arapesh (Bumbita)	APE/AON
Arapesh (Abu)	aab	Papua New Guinea	3°27S	142°57E	—	—
51.2. Urim						
Urim	uri	Papua New Guinea	3°35S	142°40E	Urim	URI
51.3. Wapei-Palei						
Au	au	Papua New Guinea	3°35S	142°05E	Au	AVT
Ningil	nin	Papua New Guinea	3°30S	142°15E	Ningil	NIZ
Olo	olo	Papua New Guinea	3°25S	142°E	Olo	ONG
Walman	wal	Papua New Guinea	3°13S	142°30E	Valman	VAN
Yil	yil	Papua New Guinea	3°32S	142°09E	Yil	YLL
51.4. West Wapei						
One	one	Papua New Guinea	3°15S	141°55E	Aunalei	AUN
52. SEPIK[27]						
52.1. Middle Sepik						
Ambulas	amb	Papua New Guinea	3°50S	143°E	Abulas	ABT
Boiken	boi	Papua New Guinea	3°30S	143°30E	Boikin	BZF
Hanga Hundi	hhu	Papua New Guinea	3°50S	142°55E	Hanga Hundi	WOS
Iatmul	iat	Papua New Guinea	4°15S	143°15E	Iatmul	IAN
Kwoma	kwo	Papua New Guinea	4°10S	142°45E	Kwoma	KMO
Manambu	mbu	Papua New Guinea	4°15S	142°50E	Manambu	MLE
52.2. Ram						
Awtuw	awt	Papua New Guinea	3°35S	142°E	Awtuw	KMN
52.3. Sepik Hill						
Alamblak	ala	Papua New Guinea	4°40S	143°20E	Alamblak	AMP
Bahinemo	bhn	Papua New Guinea	4°35S	142°50E	Bahinemo	BJH
Sare	sar	Papua New Guinea	4°30S	143°10E	Kapriman	DJU
52.4. Tama Sepik						
Mehek	mhk	Papua New Guinea	3°45S	142°30E	Mehek	NUX
Yessan-Mayo	yes	Papua New Guinea	4°10S	142°35E	Yessan-Mayo	YSS
52.5. Upper Sepik						
Abau	aba	Papua New Guinea	4°S	141°15E	Abau	AAU
Iwam	iwm	Papua New Guinea	4°20S	142°E	Iwam	IWM
Wogamusin	wog	Papua New Guinea	4°15S	142°20E	Wogamusin	WOG
52.6. Yellow River						
Namia	nam	Papua New Guinea	3°50S	141°45E	Namia	NNM
53. LOWER SEPIK-RAMU[28]						
53.1. Annaberg						
Rao	rao	Papua New Guinea	4°45S	144°30E	Rao	RAO

[26] The status of Torricelli as a family remains unclear since little information has been published for most languages in the family. *Ethnologue* also includes Monumbo, which is treated here as a language isolate.

[27] *Ethnologue* treats Sepik as part of a larger family Sepik-Ramu, which is no longer considered a family by specialists. What is treated here as Sepik corresponds to the Sepik branch of Sepik-Ramu in *Ethnologue*.

[28] *Ethnologue* treats Sepik as part of a larger family Sepik-Ramu, which is no longer considered a family by specialists. Lower Sepik-Ramu corresponds to the union of two subgroups of Sepik-Ramu in the *Ethnologue* classification, namely Ramu Proper and Nor-Pondo.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
53.2. Botin						
Kambot	kbo	Papua New Guinea	4°15S	144°08E	Ap Ma	KBX
53.3. Lower Ramu						
Kaian	kai	Papua New Guinea	4°05S	144°45E	Kaian	KCT
Mikarew	mkr	Papua New Guinea	4°25S	144°55E	Aruamu	MSY
Watam	wtm	Papua New Guinea	3°55S	144°30E	Watam	WAX
53.4. Lower Sepik[29]						
Chambri	chb	Papua New Guinea	4°10S	143°05E	Chambri	CAN
Murik	mrk	Papua New Guinea	3°50S	144°15E	Murik	MTF
Yimas	yim	Papua New Guinea	4°40S	143°33E	Yimas	YEE
53.5. Mikarew						
Kire	kie	Papua New Guinea	4°15S	144°45E	Kire	GEB
54. GAPUN[30]						
Taiap	tap	Papua New Guinea	4°05S	144°30E	Taiap	GPN
55. MONUMBO[31]						
Monumbo	mbo	Papua New Guinea	4°15S	145°E	Monumbo	MXK
56. UPPER YUAT[32]						
Haruai[33]	har	Papua New Guinea	5°05S	144°10E	Haruai	TMD
57. TRANS-NEW GUINEA[34]						
57.1. Angan						
Ampeeli	ape	Papua New Guinea	6°45S	146°05E	Ampeeli-Wojokeso	APZ
Angaatiha	ant	Papua New Guinea	7°13S	146°15E	Angaatiha	AGM
Baruya	bry	Papua New Guinea	6°55S	145°55E	Baruya	BYR
Hamtai	ham	Papua New Guinea	7°30S	146°15E	Hamtai	HMT
Menya	mey	Papua New Guinea	7°15S	146°E	Menya	MCR
57.2. Asmat-Kamoro						
Asmat	asm	Indonesia	5°30S	138°30E	5 varieties: Citak (Tamnim)/. . .	TML/. . .
Kamoro	kmr	Indonesia	4°20S	136°E	Kamoro	KGQ
57.3. Awju-Dumut						
Aghu	ahu	Indonesia	6°10S	140°10E	Aghu	AHH
Kombai	kmb	Indonesia	5°35S	140°E	Kombai	KGU
Korowai	krw	Indonesia	5°15S	140°E	2 varieties: Korowai/Korowai (North)	KHE/KRG
Pisa	pis	Indonesia	5°10S	139°10E	Awyu (Miaro)	PSA
Wambon	wbn	Indonesia	5°30S	140°25E	Wambon	WMS
57.4. Binanderean						
Binandere	bin	Papua New Guinea	8°15S	148°E	Binandere	BHG
Guhu-Samane	guh	Papua New Guinea	8°S	147°20E	Guhu-Samane	GHS
Korafe	krf	Papua New Guinea	9°03S	149°05E	Korafe	KPR
Orokaiva	oro	Papua New Guinea	8°50S	148°15E	Orokaiva	ORK
Suena	sue	Papua New Guinea	7°45S	147°33E	Suena	SUE

[29] Equivalent to Nor-Pondo in the *Ethnologue* classification.
[30] Treated as part of Sepik-Ramu by *Ethnologue*.
[31] Treated by *Ethnologue* as belonging to Torricelli.
[32] Treated by *Ethnologue* as a subgroup of Ramu within Sepik-Ramu.
[33] Treated by *Ethnologue* as belonging to Trans-New Guinea.
[34] The version of Trans-New Guinea here is considerably smaller than that in *Ethnologue*. The following groups are treated here as separate families but as branches of Trans-New Guinea in *Ethnologue:* West Bomberai, Kwerba, Tor, Nimboran, Sentani, Morwap, Kaure, Border, Senagi, Oksapmin, Kayagar, Marind, Kolopom, Mombum, Bulaka River, Morehead and Upper Maro Rivers, Awin-Pare, East Strickland, Bosavi, Gogodala-Suki, Kiwaian, Turama-Kikorian, Teberan-Pawaian, Eleman, Yareban, and Dagan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
57.5. Chimbu						
Chuave	cve	Papua New Guinea	6°07S	145°07E	Chuave	CJV
Golin	gln	Papua New Guinea	6°20S	144°45E	Golin	GVF
Kaugel	kgl	Papua New Guinea	6°S	144°E	Umbu-Ungu	UMB
Kuman	kmn	Papua New Guinea	5°55S	145°E	Kuman	KUE
Maring	mav	Papua New Guinea	5°30S	144°40E	Maring	MBW
Salt-Yui	syu	Papua New Guinea	6°17S	145°02E	Salt-Yui	SLL
Tabare	tba	Papua New Guinea	6°05S	145°E	dialect of Sinasina	dSST
Wahgi	wah	Papua New Guinea	5°50S	144°43E	2 varieties: Wahgi/Wahgi (North)	WAK/WHG
57.6. Dani						
Dani (Lower Grand Valley)	dni	Indonesia	4°20S	138°50E	Dani (Lower Grand Valley)	DNI
Yali	yli	Indonesia	4°S	139°20E	3 varieties: Yali (Angguruk)/. . .	YLI/. . .
57.7. Duna-Bogaya						
Duna	dun	Papua New Guinea	5°30S	142°30E	Duna	DUC
57.8. Eastern Highlands						
Agarabi	aga	Papua New Guinea	6°10S	146°E	Agarabi	AGD
Auyana	auy	Papua New Guinea	6°30S	145°45E	Awiyaana	AUY
Awa	awa	Papua New Guinea	6°40S	145°45E	Awa	AWB
Benabena	beb	Papua New Guinea	6°10S	145°30E	Benabena	BEF
Binumarien	bnm	Papua New Guinea	6°17S	146°05E	Binumarien	BJR
Fore	for	Papua New Guinea	6°45S	145°30E	Fore	FOR
Gadsup	gds	Papua New Guinea	6°15S	146°E	Gadsup	GAJ
Gahuku	gah	Papua New Guinea	6°S	145°25E	Alekano	GAH
Hua	hua	Papua New Guinea	6°20S	145°20E	dialect of Yagaria	dYGR
Kamano-Kafe	kak	Papua New Guinea	6°15S	145°40E	Kamano	KBQ
Siane	sia	Papua New Guinea	6°05S	145°12E	Siane	SNP
Tairora	trr	Papua New Guinea	6°30S	146°E	Tairora	TBG
Usarufa	usr	Papua New Guinea	6°25S	145°35E	Usarufa	USA
Yagaria	ygr	Papua New Guinea	6°20S	145°25E	Yagaria	YGR
57.9. Engan						
Enga	ena	Papua New Guinea	5°30S	143°40E	Enga	ENQ
Kewa	kew	Papua New Guinea	6°30S	143°50E	2 varieties: Kewa (East)/Kewa (West)	KJS/KEW
Kyaka	kyk	Papua New Guinea	5°33S	144°05E	Kyaka	KYC
57.10. Finisterre-Huon						
Burum	bmr	Papua New Guinea	6°30S	147°20E	Burum-Mindik	BMU
Kâte	kat	Papua New Guinea	6°30S	147°45E	Kâte	KMG
Nabak	nab	Papua New Guinea	6°25S	147°E	Nabak	NAF
Nankina	nnk	Papua New Guinea	5°47S	146°27E	Nankina	NNK
Ono	ono	Papua New Guinea	6°10S	147°35E	Ono	ONS
Rawa	rwa	Papua New Guinea	5°50S	146°E	Rawa	RWO
Selepet	slp	Papua New Guinea	6°10S	147°10E	Selepet	SEL
Wantoat	wnt	Papua New Guinea	6°10S	146°30E	Wantoat	WNC
57.11. Goilalan						
Kunimaipa	kmp	Papua New Guinea	8°S	146°50E	Kunimaipa	KUP
Weri	wer	Papua New Guinea	7°45S	146°55E	Weri	WER
57.12. Koiarian						
Barai	baa	Papua New Guinea	9°20S	148°05E	Barai	BCA
Koiali (Mountain)	kmo	Papua New Guinea	9°S	147°30E	Koiali (Mountain)	KPX
Koiari	koi	Papua New Guinea	9°30S	147°20E	Koiari (Grass)	KBK
Koita	kta	Papua New Guinea	9°20S	147°05E	Koitabu	KQI
Ömie	omi	Papua New Guinea	9°03S	148°05E	Ömie	AOM
57.13. Kutubuan						
Fasu	fas	Papua New Guinea	6°35S	143°20E	Fasu	FAA
Foe	foe	Papua New Guinea	6°30S	143°30E	Foi	FOI
57.14. Madang						
Amele	ame	Papua New Guinea	5°15S	145°35E	Amele	AMI
Anamuxra	amx	Papua New Guinea	4°40S	145°E	Anamgura	IMI

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Bau	bau	Papua New Guinea	5°15S	145°37E	Bau	BBD
Bongu	bon	Papua New Guinea	5°30S	145°55E	Bongu	BPU
Garus	grs	Papua New Guinea	4°57S	145°40E	Garus	GYB
Kalam	klq	Papua New Guinea	5°15S	144°35E	Kalam	KMH
Kobon	kob	Papua New Guinea	5°10S	144°20E	Kobon	KPW
Mugil	mug	Papua New Guinea	4°53S	145°45E	Bargam	MLP
Nend	nnd	Papua New Guinea	5°S	144°55E	Nend	ANH
Siroi	sro	Papua New Guinea	5°30S	146°E	Siroi	SSD
Tauya	tau	Papua New Guinea	5°45S	145°20E	Tauya	TYA
Usan	usa	Papua New Guinea	4°50S	145°10E	Usan	WNU
Waskia	wsk	Papua New Guinea	4°30S	146°E	Waskia	WSK
57.15. Mailuan						
Magi	mgi	Papua New Guinea	10°20S	149°20E	Mailu	MGU
57.16. Mairasi-Tanah Merah						
Mairasi	mrs	Indonesia	3°30S	134°E	Mairasi	FRY
57.17. Mek						
Eipo	eip	Indonesia	4°20S	140°05E	Eipomek	EIP
Ketengban	ktn	Indonesia	4°30S	140°30E	Ketengban	KIN
Una	una	Indonesia	4°40S	140°E	Una	MTG
Yale (Kosarek)	yal	Indonesia	4°05S	139°30E	Yale (Kosarek)	KKL
57.18. Ok						
Kati (in West Papua, Indonesia)	kti	Indonesia	5°45S	140°55E	dialect of Yongkom	dYON
Telefol	tlf	Papua New Guinea	5°S	141°45E	Telefol	TLF
Tifal	tif	Papua New Guinea	5°S	141°20E	Tifal	TIF
57.19. Timor-Alor-Pantar						
Kolana	kln	Indonesia	8°15S	125°07E	Wersing	KVW
Kui (in Indonesia)	kiu	Indonesia	8°26S	124°33E	Kui	KVD
Tanglapui	tgp	Indonesia	8°17S	125°05E	Sawila	SWT
Woisika	woi	Indonesia	8°15S	124°50E	Kamang	WOI
57.20. Wissel Lakes-Kemandoga						
Ekari	eka	Indonesia	3°50S	135°30E	Ekari	EKG
Moni	mni	Indonesia	3°40S	137°E	Moni	MNZ
58. KAYAGAR[35]						
Tamagario	tmg	Indonesia	6°25S	139°15E	Tamagario	TCG
59. MARIND[36]						
59.1. Marind Proper						
Boazi	boa	Papua New Guinea	7°S	141°20E	Boazi	KVG
Marind	mrd	Indonesia	7°50S	140°10E	Marind	MRZ
Yaqay	yqy	Indonesia	6°35S	139°15E	Yaqay	JAQ
Zimakani	zim	Papua New Guinea	7°26S	141°26E	Zimakani	ZIK
59.2. South Bird's Head						
Arandai	ard	Indonesia	2°05S	133°E	Arandai	JBJ
Inanwatan	ina	Indonesia	2°05S	132°05E	Suabo	SZP
60. KOLOPOM[37]						
Kimaghama	kim	Indonesia	8°S	138°30E	Kimaama	KIG
Riantana	ria	Indonesia	7°30S	138°30E	Riantana	RAN

[35] Treated by *Ethnologue* as a branch of Trans–New Guinea.

[36] The two genera within the Marind family are both treated as belonging to Trans–New Guinea in the Ethnologue classification, but are not treated as closely related to each other within Trans–New Guinea, so that there is no group corresponding to the Marind family in the Ethnologue classification.

[37] Treated by *Ethnologue* as a branch of Trans–New Guinea.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
61. MOMBUM[38]						
Mombum	mom	Indonesia	8°15S	138°45E	Mombum	MSO
62. BULAKA RIVER[39]						
Maklew	mkl	Indonesia	7°40S	139°25E	Maklew	MGF
Yelmek	ylm	Indonesia	7°40S	139°10E	Yelmek	JEL
63. MOREHEAD AND UPPER MARO RIVERS[40]						
Kanum (Bädi)	knb	Indonesia	8°50S	140°45E	Kanum (Bädi)	KHD
Kanum (Ngkâlmpw)	knp	Indonesia	8°40S	140°55E	Kanum (Ngkâlmpw)	KCD
Moraori	mri	Indonesia	8°35S	140°40E	Morori	MOK
Yei	yei	Papua New Guinea, Indonesia	7°55S	140°55E	Yei	JEI
64. AWIN-PARE[41]						
Awin	awi	Papua New Guinea	5°45S	141°25E	Aekyom	AWI
65. EAST STRICKLAND[42]						
Nomad	nmd	Papua New Guinea	6°10S	142°15E	Samo	SMQ
66. BOSAVI[43]						
Edolo	edo	Papua New Guinea	6°10S	142°40E	Edolo	ETR
Kaluli	kll	Papua New Guinea	6°30S	142°45E	Kaluli	BCO
67. GOGODALA-SUKI[44]						
Gogodala	gog	Papua New Guinea	8°05S	142°50E	Gogodala	GOH
Suki	suk	Papua New Guinea	8°S	141°45E	Suki	SUI
68. KIWAIAN[45]						
Kiwai	kiw	Papua New Guinea	8°S	143°30E	2 varieties: Kiwai (Northeast)/ Kiwai (Southern)	KIW/KJD
69. WESTERN FLY						
Meryam Mir	mer	Australia	9°55S	144°05E	Meriam	ULK
70. TURAMA-KIKORIAN[46]						
Rumu	rum	Papua New Guinea	7°10S	144°15E	Rumu	KLQ
71. TEBERAN-PAWAIAN[47]						
71.1. Pawaian						
Pawaian	paw	Papua New Guinea	7°00S	145°05E	Pawaia	PWA
71.2. Teberan						
Dadibi	dad	Papua New Guinea	6°33S	144°35E	Dadibi	MPS
Podopa	pdp	Papua New Guinea	7°S	144°30E	Folopa	PPO

[38] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[39] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[40] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[41] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[42] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[43] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[44] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[45] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[46] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[47] Treated by *Ethnologue* as a branch of Trans-New Guinea.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
72. ELEMAN[48]						
Kaki Ae	kae	Papua New Guinea	8°S	145°50E	Kaki Ae	TBD
Orokolo	orl	Papua New Guinea	7°50S	145°20E	Orokolo	ORO
Toaripi	toa	Papua New Guinea	8°20S	146°15E	Toaripi	TPI
73. YAREBAN[49]						
Yareba	yar	Papua New Guinea	9°30S	148°30E	Yareba	YRB
74. DAGAN[50]						
Daga	dag	Papua New Guinea	10°S	149°20E	Daga	DGZ
75. YELE[51]						
Yelî Dnye	yel	Papua New Guinea	11°22S	154°10E	Yele	YLE
76. ANEM[52]						
Anem	ane	Papua New Guinea	5°35S	149°05E	Anem	ANZ
77. SULKA[53]						
Sulka	sul	Papua New Guinea	5°05S	151°55E	Sulka	SLK
78. BAINING-TAULIL[54]						
Baining	bng	Papua New Guinea	4°35S	152°E	Qaqet	BYX
Mali	mli	Papua New Guinea	4°50S	152°15E	Mali	GCC
Taulil	tll	Papua New Guinea	4°25S	152°05E	dialect of Taulil-Butam	dTUH
79. KUOT[55]						
Kuot	kuo	Papua New Guinea	3°03S	151°30E	Kuot	KTO
80. WEST BOUGAINVILLE[56]						
Konua	knu	Papua New Guinea	5°50S	154°50E	Rapoisi	KYX
Rotokas	rtk	Papua New Guinea	6°S	155°10E	Rotokas	ROO
81. EAST BOUGAINVILLE[57]						
Buin	bun	Papua New Guinea	6°45S	155°45E	Buin	BUO
Motuna	mot	Papua New Guinea	6°37S	155°25E	Siwai	SIW
Nasioi	nas	Papua New Guinea	6°20S	155°35E	Naasioi	NAS
82. SOLOMONS EAST PAPUAN[58]						
Bilua	bil	Solomon Islands	7°45S	156°40E	Bilua	BLB
Lavukaleve	lav	Solomon Islands	9°05S	159°12E	Lavukaleve	LVK
Savosavo	svs	Solomon Islands	9°08S	159°48E	Savosavo	SVS

[48] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[49] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[50] Treated by *Ethnologue* as a branch of Trans-New Guinea.
[51] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[52] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[53] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[54] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[55] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[56] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[57] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.
[58] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
83. *REEF ISLANDS—SANTA CRUZ*[59]						
Ayiwo	ayw	Solomon Islands	10°20S	166°15E	Ayiwo	NFL
Nambakaengö	nbk	Solomon Islands	10°47S	165°52E	Santa Cruz	STC
84. AUSTRALIAN[60]						
84.1. Bunuban						
Bunuba	bnb	Australia	17°45S	125°45E	Bunaba	BCK
Gooniyandi	goo	Australia	18°20S	126°20E	Gooniyandi	GNI
84.2. Burarran						
Burarra	bua	Australia	12°15S	134°35E	Burarra	BVR
Gurr-goni	grg	Australia	12°15S	134°25E	Guragone	GGE
84.3. Daly[61]						
84.3.1. Anson Bay						
Bachamal	bac	Australia	13°10S	130°10E	Wadjiginy	WDJ
Pungupungu	pun	Australia	13°30S	130°25E	Kuwama	QKU
84.3.2. Eastern Daly						
Kamu	kmw	Australia	13°35S	130°50E	Kamu	QKY
Madngele	mdl	Australia	13°50S	130°25E	Madngele	ZML
84.3.3. Murrinh-Patha						
Murrinh-Patha	mpa	Australia	14°40S	129°40E	Murrinh-Patha	MWF
84.3.4. Northern Daly						
Malakmalak	mlk	Australia	13°25S	130°25E	Mullukmulluk	MPB
Tyeraity	tye	Australia	13°S	130°20E	Tyaraity	WOA
84.3.5. Southern Daly						
Ngan'gityemerri	ngg	Australia	14°15S	130°25E	dialect of Nangikurrunggurr	dNAM
Ngankikurungkurr	ngk	Australia	14°S	130°30E	Nangikurrunggurr	NAM
84.3.6. Western Daly						
Emmi	emm	Australia	13°30S	130°E	Ami	AMY
Maranungku	mku	Australia	13°40S	130°E	Maranunggu	ZMR
Maringarr	mrr	Australia	14°15S	130°E	Maringarr	ZMT
Marrithiyel	mrh	Australia	13°50S	130°E	Marithiel	MFR
84.4. Djeragan						
Kitja	kij	Australia	17°30S	127°45E	Kitja	GIA
Miriwung	mir	Australia	15°40S	129°E	Miriwung	MEP
84.5. Gaagudju						
Gaagudju	gaa	Australia	12°35S	132°50E	Gagadu	GBU
84.6. Garrwan						
Garrwa	grr	Australia	17°05S	137°10E	Garawa	GBC
84.7. Gunwinyguan						
84.7.1. Anindilyakwa						
Anindilyakwa	ann	Australia	14°S	136°30E	Anindilyakwa	AOI

[59] Treated by *Ethnologue* as a branch of East Papuan, which is not recognized here as a family.

[60] The claim that all of the indigenous languages of Australia (other than the Tasmanian languages and Meriam, a language spoken on Murray Island in the Torres Straits between Australia and Papua New Guinea and belong to the Western Fly family, whose other languages are spoken in Papua New Guinea) remains a hypothesis, albeit one assumed by many specialists. For some of the groups placed in this family, the evidence is actually fairly weak. However, because we are not sure just which groups are those for which the evidence is weak, we follow *Ethnologue* in treating all of these languages as belonging to the Australian family.

[61] The subclassification of Daly is partly due to suggestions by Evans (2003), citing work in progress by Ian Green, Mark Harvey, and Nick Reid.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
84.7.2. Gunwinygic						
Bininj Gun-Wok	bbw	Australia	12°30S	133°45E	Gunwinggu	GUP
Gunbalang	gnb	Australia	12°S	134°E	Kunbarlang	WLG
Ngalkbun	nkb	Australia	13°30S	134°50E	Ngalkbun	NGK
84.7.3. Ngalakan						
Ngalakan	ngl	Australia	14°15S	134°E	Ngalakan	NIG
84.7.4. Ngandi						
Ngandi	ngn	Australia	13°50S	135°E	Ngandi	NID
84.7.5. Nunggubuyu						
Nunggubuyu	nug	Australia	13°45S	135°40E	Nunggubuyu	NUY
84.7.6. Rembarnga						
Rembarnga	rmb	Australia	12°50S	134°35E	Rembarunga	RMB
84.7.7. Waray						
Waray (in Australia)	wry	Australia	13°10S	131°15E	Waray	WRZ
84.7.8. Yangmanic						
Wagiman	wag	Australia	14°S	131°15E	Wageman	WAQ
Wardaman	wrd	Australia	15°30S	131°E	Wardaman	WRR
84.8. Iwaidjan						
Iwaidja	iwa	Australia	11°30S	132°40E	Iwaidja	IBD
Maung	mau	Australia	11°55S	133°30E	Maung	MPH
84.9. Jaminjungan						
Jaminjung	jam	Australia	15°05S	130°30E	Djamindjung	DJD
84.10. Laragiyan						
Laragia	lar	Australia	12°40S	130°50E	Laragia	LRG
84.11. Limilngan						
Limilngan	lml	Australia	12°30S	131°55E	Limilngan	LMC
84.12. Mangarrayi[62]						
Mangarrayi	myi	Australia	14°40S	133°30E	Mangarayi	MPC
84.13. Maran						
Alawa	alw	Australia	15°10S	134°15E	Alawa	ALH
Mara	mra	Australia	15°S	135°10E	Mara	MEC
Warndarang	wrn	Australia	14°25S	135°35E	Wandarang	WND
84.14. Nakkara[63]						
Nakkara	nkk	Australia	12°7S	134°25E	Nakara	NCK
84.15. Ndjébbana[64]						
Ndjébbana	ndj	Australia	12°10S	134°07E	Djeebbana	DJJ
84.16. Nyulnyulan						
Bardi	brd	Australia	16°35S	122°55E	Baadi	BCJ
Nyigina	nyn	Australia	18°S	124°20E	Nyigina	NYH
Nyulnyul	nyu	Australia	17°S	122°50E	Nyulnyul	NYV
Warrwa	wrw	Australia	17°30S	123°30E	—	—
Yawuru	ywr	Australia	18°S	122°30E	Yawuru	YWR

[62] Treated by *Ethnologue* as belonging to the Gunwinyguan subfamily.
[63] Treated by *Ethnologue* as belonging to Burarran.
[64] Treated by *Ethnologue* as belonging to Burarran.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
84.17. Pama-Nyungan						
Adynyamathanha	adn	Australia	30°30S	139°30E	Adynyamathanha	ADT
Alyawarra	aly	Australia	22°30S	135°E	Alyawarr	ALY
Anguthimri	agt	Australia	12°20S	141°50E	—	—
Arabana	abn	Australia	28°15S	136°15E	Arabana	ARD
Arrernte	arr	Australia	24°S	134°E	2 varieties: Arrernte (Eastern)/ Arrarnta (Western)	AER/ARE
Arrernte (Mparntwe)	amp	Australia	24°S	136°E	Arrernte (Eastern)	AER
Arrernte (Western)	awe	Australia	24°S	132°30E	Arrarnta (Western)	ARE
Badimaya	bdm	Australia	27°40S	118°E	Badimaya	BIA
Bandjalang	bnj	Australia	28°45S	153°E	Bandjalang	BDY
Bandjalang (Casino)	bca	Australia	28°55S	153°E	dialect of Bandjalang	dBDY
Bandjalang (Waalubal)	bwa	Australia	29°05S	152°35E	dialect of Bandjalang	dBDY
Bandjalang (Yugumbir)	byu	Australia	27°55S	153°E	dialect of Bandjalang	dBDY
Banggarla	bnl	Australia	32°S	137°E	Banggarla	BJB
Bilinara	bnr	Australia	16°S	131°40E	dialect of Ngarinman	dNBJ
Biri	bii	Australia	20°30S	146°30E	Biri	BZR
Bularnu	bnu	Australia	21°S	138°E	—	—
Colac	clc	Australia	38°20S	143°30E	—	—
Dharawal	dhw	Australia	34°30S	150°30E	Thurawal	TBH
Dhargari	dhr	Australia	23°45S	114°55E	Dhargari	DHR
Dharumbal	dhb	Australia	22°45S	150°30E	dialect of Bayali	dBJY
Dhurga	dhu	Australia	35°40S	150°E	Dhurga	DHU
Dhuwal (Dätiwuy)	dda	Australia	12°10S	136°15E	dialect of Dhuwal	dDUJ
Diyari	diy	Australia	28°S	139°E	Dieri	DIF
Djabugay	dja	Australia	16°45S	145°35E	Dyaabugay	DYY
Djambarrpuyngu	djm	Australia	12°10S	135°30E	Djambarrpuyngu	DJR
Djapu	djp	Australia	12°40S	136°E	dialect of Dhuwal	dDUJ
Djaru	djr	Australia	18°45S	128°E	Jaru	DDJ
Djinang	djn	Australia	12°20S	134°50E	Djinang	DJI
Dyirbal	dyi	Australia	17°50S	145°35E	Dyirbal	DBL
Gaalpu	glp	Australia	12°40S	136°45E	dialect of Dhangu	dGLA
Gamilaraay	gml	Australia	29°50S	149°30E	Kamilaroi	KLD
Gidabal	gid	Australia	28°25S	152°25E	dialect of Bandjalang	dBDY
Gugada	gug	Australia	30°S	134°E	Kokata	KTD
Gumatj	gmt	Australia	12°30S	135°30E	Gumatj	GNN
Gumbaynggir	gum	Australia	30°10S	152°30E	Kumbainggar	KGS
Gunya	gny	Australia	26°30S	146°30E	Gunya	GYY
Gureng Gureng	ggu	Australia	25°15S	151°E	Gureng Gureng	GNR
Gurindji	gji	Australia	17°40S	130°40E	Gurinji	GUE
Guugu Yimidhirr	guu	Australia	15°S	144°50E	Guguyimidjir	KKY
Juat	jua	Australia	31°S	115°30E	—	—
Kala Lagaw Ya	kly	Australia	10°07S	142°07E	Kala Lagaw Ya	MWP
Kalkatungu	kgu	Australia	21°S	139°30E	Kalkutung	KTG
Karadjeri	krj	Australia	19°S	122°E	Karadjeri	GBD
Kaurna	kaq	Australia	34°15S	138°30E	—	—
Kugu Nganhcara	knc	Australia	14°25S	142°E	Kuku-Uwanh	UWA
Kuku-Yalanji	kya	Australia	16°S	145°E	Kuku-Yalanji	GVN
Kunjen	kjn	Australia	16°30S	142°30E	Kunjen	KJN
Kuuku Ya'u	kuu	Australia	12°35S	143°05E	Kuuku-Ya'u	QKL
Lamu-Lamu	lla	Australia	14°50S	144°30E	Lamu-Lamu	LBY
Lardil	lrd	Australia	16°30S	139°20E	Lardil	LBZ
Madimadi	mdm	Australia	34°30S	143°30E	—	—
Mantjiltjara	mnj	Australia	22°40S	125°05E	dialect of Martu Wangka	dMPJ
Margany	mny	Australia	27°S	144°30E	Margany	ZMC
Martu Wangka	mwa	Australia	21°30S	126°E	Martu Wangka	MPJ
Martuthunira	mrt	Australia	20°50S	116°30E	Martuyhunira	VMA
Mayi-Yapi	myy	Australia	20°S	141°E	—	—
Mbabaram	mbb	Australia	17°10S	145°E	—	—
Mudburra	mdb	Australia	17°20S	132°05E	Mudbura	MWD
Muruwari	mrw	Australia	28°45S	147°E	Muruwari	ZMU
Ngaanyatjarra	ntj	Australia	26°S	126°30E	Ngaanyatjarra	NTJ
Ngadjumaja	ngj	Australia	32°20S	123°50E	Ngadjunmaya	NJU
Ngarinyeri	ngr	Australia	36°S	140°E	Narrinyeri	NAY
Ngarinyman	ngy	Australia	16°30S	130°30E	Ngarinman	NBJ
Ngarluma	nlu	Australia	20°50S	117°30E	Ngarluma	NRL
Ngawun	ngw	Australia	19°S	141°30E	Ngawun	NXN
Ngiyambaa	ngi	Australia	31°45S	145°30E	dialect of Wangaaybuwan-Ngiyambaa	dWYB
Nhanda	nha	Australia	27°S	114°10E	—	—
Nyangumarda	nyr	Australia	20°S	121°E	Nyangumarta	MMA
Nyawaygi	nya	Australia	19°S	146°10E	Nyawaygi	NYT
Nyungar	nju	Australia	33°45S	122°E	Nyunga	NYS
Paakantyi	pkn	Australia	32°30S	142°30E	Darling	DRL

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Pakanha	pak	Australia	14°30S	142°25E	Pakanha	PKN
Panyjima	pny	Australia	22°50S	118°25E	Panytyima	PNW
Pintupi	pin	Australia	23°S	129°E	Pintupi-Luritja	PIU
Pitjantjatjara	pit	Australia	26°S	130°E	Pitjantjatjara	PJT
Pitta Pitta	ppi	Australia	22°50S	140°E	Pitta Pitta	PIT
Ritharngu	rit	Australia	12°50S	135°30E	Ritarungo	RIT
Thaayorre	thy	Australia	14°35S	141°50E	Thayore	THD
Thaypan	thp	Australia	14°50S	143°10E	Thaypan	TYP
Uradhi	uhi	Australia	11°55S	142°25E	Uradhi	URF
Walmatjari	wlm	Australia	19°30S	125°45E	Walmajarri	WMT
Wangkumara	wan	Australia	27°20S	143°E	dialect of Ngura	dNBX
Warlpiri	wrl	Australia	20°S	132°20E	Warlpiri	WBP
Warluwara	wlw	Australia	22°S	138°30E	Warluwara	WRB
Warrgamay	wrg	Australia	18°35S	146°05E	Warrgamay	WGY
Warrnambool	wrb	Australia	38°15S	142°30E	—	—
Warrungu	wgu	Australia	18°S	145°E	Warungu	WRG
Wathawurrung	wth	Australia	38°S	144°E	—	—
Watjarri	wat	Australia	26°S	117°30E	Wajarri	WBV
Wembawemba	wem	Australia	35°20S	144°E	—	—
Western Desert (Ooldea)	wdo	Australia	30°30S	132°E	—	—
Wik Munkan	wmu	Australia	13°55S	141°45E	Wik-Mungkan	WIM
Wik Ngathana	wn	Australia	13°55S	141°30E	Wik-Ngathana	WIG
Wirangu	wir	Australia	32°S	134°E	Wirangu	WIW
Woiwurrung	wwr	Australia	37°30S	145°30E	—	—
Yankuntjatjara	ynk	Australia	27°S	132°E	Yankunytjatjara	KDD
Yanyuwa	yny	Australia	16°25S	137°10E	Yanyuwa	JAO
Yaygir	yyg	Australia	29°40S	153°E	—	—
Yidiny	yid	Australia	17°S	145°45E	Yidiny	YII
Yindjibarndi	yin	Australia	21°45S	117°55E	Yindjibarndi	YIJ
Yingkarta	yng	Australia	25°10S	114°50E	Yinggarda	YIA
Yir Yiront	yir	Australia	14°50S	141°50E	Yir Yiront	YIY
Yorta Yorta	yyo	Australia	36°S	146°E	—	—
Yulparija	ylb	Australia	21°S	124°E	dialect of Martu Wangka	dMPJ
Yuwaalaraay	yuw	Australia	29°30S	148°E	dialect of Kamilaroi	dKLD
84.18. Tangkic[65]						
Kayardild	kay	Australia	17°03S	139°30E	Gayardilt	GYD
Yukulta	yuk	Australia	17°20S	138°50E	Ganggalida	GCD
84.19. Tiwian						
Tiwi	tiw	Australia	11°30S	131°E	Tiwi	TIW
84.20. West Barkly						
Djingili	dji	Australia	17°45S	134°E	Djingili	JIG
Wambaya	wam	Australia	18°40S	135°45E	Wambaya	WMB
84.21. Wororan						
Gunin	gnn	Australia	14°15S	126°40E	Kwini	GWW
Ungarinjin	ung	Australia	16°20S	126°E	Ngarinyin	UNG
Worora	wor	Australia	15°40S	124°40E	Worora	UNP
85. TASMANIAN						
Tasmanian[66]	tsm	Australia	42°10S	146°30E	—	—
Tasmanian (Oyster Bay to Pitwater)	toy	Australia	42°30S	147°45E	—	—
86. ESKIMO-ALEUT						
Aleut	ale	United States	54°N	166°W	Aleut	ALW
Aleut (Eastern)	aea	United States	54°45N	164°W	dialect of Aleut	dALW
Alutiiq	atq	United States	57°N	157°W	Yupik (Pacific Gulf)	EMS
Greenlandic (East)	gre	Greenland	65°N	40°W	dialect of Inuktitut (Greenlandic)	dESG
Greenlandic (South)	gso	Greenland	60°N	44°W	dialect of Inuktitut (Greenlandic)	dESG

[65] Treated by *Ethnologue* as a branch of Pama-Nyungan.

[66] Tasmanian is actually a group of languages; however, some sources do not make clear which language is being described or imply that the properties are true of Tasmanian languages in general.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Greenlandic (West)	grw	Greenland	64°N	51°W	dialect of Inuktitut (Greenlandic)	dESG
Inuktitut (Quebec-Labrador)	iql	Canada	62°N	73°W	dialect of Inukititut (Eastern Canadian)	dESB
Inuktitut (Rankin Inlet)	inr	Canada	63°N	92°W	dialect of Inukititut (Western Canadian)	dESC
Inuktitut (Salluit)	ins	Canada	62°N	76°W	dialect of Inukititut (Eastern Canadian)	dESB
Iñupiaq	inu	United States	67°N	161°W	Inupiatun (North Alaskan)	ESI
Kangiryuarmiut	kgt	Canada	68°20N	133°45W	dialect of Inukititut (Western Canadian)	dESC
Yup'ik (Central)	ypk	United States	59°30N	160°W	Yupik (Central)	ESU
Yup'ik (Chevak)	ych	United States	61°30N	165°45W	dialect of Yupik (Central)	dESU
Yupik (Naukan)	yna	Russia	66°N	172°W	Yupik (Naukan)	YNK
Yup'ik (Norton Sound)	yun	United States	64°N	161°W	dialect of Yup'ik (Central)	dESU
Yupik (Siberian)	yus	Russia	65°N	173°W	3 varieties: Yupik (Central Siberian)/. . .	ESS/. . .
Yupik (Sirenik)	ysi	Russia	64°30N	174°W	Yupik (Sirenik)	YSR
Yupik (St. Lawrence Island)	ysl	United States	63°30N	170°30W	dialect of Yupik (Central Siberian)	dESS

87. NA-DENE[67]

87.1. Athapaskan

Ahtna	aht	United States	62°N	145°W	Ahtena	AHT
Apache (Chiricahua)	apc	United States	33°15N	108°W	Apache (Mescalero-Chiricahua)	APM
Apache (Jicarilla)	apj	United States	36°35N	104°W	Apache (Jicarilla)	APJ
Apache (Western)	apw	United States	33°45N	110°W	Apache (Western)	APW
Carrier	crq	Canada	53°45N	123°30W	Carrier	CAR
Chasta Costa	cco	United States	42°40N	124°W	—	—
Chipewyan	chp	Canada	59°N	106°W	Chipewyan	CPW
Degexit'an	dgx	United States	62°N	160°W	Degexit'an	ING
Hupa	hup	United States	41°05N	123°40W	Hupa	HUP
Kato	kto	United States	39°40N	123°40W	Kato	KTW
Koyukon	kyn	United States	65°N	155°W	Koyukon	KOY
Kutchin	kth	United States, Canada	67°N	146°W	Gwich'in	KUC
Mattole	mtl	United States	40°10N	124°10W	Mattole	MVB
Navajo	nav	United States	36°10N	108°W	Navajo	NAV
Sarcee	src	Canada	52°30N	116°W	Sarsi	SRS
Slave	sla	Canada	67°N	125°W	2 varieties: Slavey (North)/Slavey (South)	SCS/SLA
Slavey	slv	Canada	60°N	120°W	dialect of Slavey (Sorth)	dSLA
Tanacross	tnc	United States	63°30N	143°30W	Tanacross	TCB
Tanaina	tnj	United States	62°N	150°W	Tanaina	TFN
Tanana (Lower)	tnl	United States	65°N	150°W	Tanana (Lower)	TAA
Tutchone (Northern)	tut	Canada	63°N	137°W	Tutchone (Northern)	TUT
Upper Kuskokwim	uku	United States	63°N	157°W	Kuskokwim (Upper)	KUU

87.2. Eyak

Eyak	eya	United States	60°30N	145°W	Eyak	EYA

87.3. Tlingit

Tlingit	tli	United States	59°N	135°W	Tlingit	TLI

88. HAIDA[68]

Haida	hai	Canada, United States	53°N	132°W	2 varieties: Haida (Northern)/Haida (Southern)	HAI/HAX
Haida (Northern)	hno	Canada	54°N	132°30W	Haida (Northern)	HAI

89. ALGIC

89.1. Algonquian

Abenaki (Western)	abw	United States, Canada	44°N	72°15W	Abnaki (Western)	ABE
Algonquin	alg	Canada	46°N	77°W	Algonquin	ALG
Arapaho	aho	United States	40°N	103°W	Arapaho	ARP
Atikamekw	ati	Canada	48°N	74°W	Atikamekw	TET

[67] *Ethnologue* also includes Haida, which is treated here as a language isolate.

[68] Treated by *Ethnologue* as belonging to Na-Dene.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Blackfoot	bla	Canada, United States	50°N	112°40W	Blackfoot	BLC
Cheyenne	cyn	United States	43°30N	104°W	Cheyenne	CHY
Chippewa (Red Lake and Pillager)	cpw	United States	48°N	95°W	dialect of Chippewa	dCIW
Cree (Eastern)	cea	Canada	51°N	74°W	2 varieties: Cree (Northern East)/ Cree (Southern East)	CRL/CRE
Cree (Plains)	cre	Canada	54°N	110°W	Cree (Plains)	CRP
Fox	fox	United States	43°N	83°W	Mesquakie	SAC
Illinois	ill	United States	40°N	90°W	Miami	MIA
Kickapoo	kic	United States	42°15N	84°W	Kickapoo	KIC
Mahican	mhc	United States	42°40N	73°30W	dialect of Mohegan-Montauk-Narragansett	dMOF
Massachusett	mtt	United States	42°30N	71°W	Wampanoag	WAM
Menomini	men	United States	45°30N	88°20W	Menomini	MEZ
Michif	mcf	Michif	52°N	100°10W	Michif	CRG
Micmac	mic	Canada	45°N	63°W	Micmac	MIC
Montagnais	mtg	Canada	52°N	65°W	Montagnais	MOE
Munsee	mse	United States	41°30N	74°40W	Munsee	UMU
Nanticoke	nnt	United States	38°N	76°15W	Nanticoke	NNT
Naskapi	nsk	Canada	56°N	70°W	Naskapi	NSK
Ojibwa (Eastern)	oji	Canada	46°N	80°W	Ojibwa (Eastern)	OJG
Ojibwa (Severn)	ojs	Canada	53°N	90°W	Ojibwa (Severn)	OJS
Ojibwe (Minnesota)	ojm	United States	47°40N	92°30W	dialect of Chippewa	dCIW
Passamaquoddy-Maliseet	psm	United States, Canada	45°N	67°W	Malecite-Passamaquoddy	MAC
Potawatomi	pot	United States	44°30N	85°W	Potawatomi	POT
Powhatan	pow	United States	37°20N	76°30W	Powhatan	PIM
Shawnee	shw	United States	40°N	83°W	Shawnee	SJW
Unami	unm	United States	40°N	75°10W	Unami	DEL
89.2. Wiyot						
Wiyot	wiy	United States	40°50N	124°10W	Wiyot	WIY
89.3. Yurok						
Yurok	yur	United States	41°20N	124°W	Yurok	YUR
90. BEOTHUK						
Beothuk	beo	Canada	48°N	57°W	Beothuk	BUE
91. IROQUOIAN						
91.1. Northern Iroquoian						
Cayuga	cyg	United States	42°45N	76°45W	Cayuga	CAY
Mohawk	moh	United States, Canada	43°30N	74°15W	Mohawk	MOH
Oneida	ond	United States	43°N	75°40W	Oneida	ONE
Onondaga	onn	United States	42°45N	76°10W	Onondaga	ONO
Seneca	snc	United States	42°30N	77°30W	Seneca	SEE
Tuscarora	tus	United States	36°N	77°30W	Tuscarora	TUS
Wyandot	wya	Canada	44°20N	77°30W	Wyandot	WYA
91.2. Southern Iroquoian						
Cherokee	che	United States	35°30N	83°30W	Cherokee	CER
92. TIMUCUA						
Timucua	tmc	United States	30°15N	82°30W	—	—
93. YUCHI						
Yuchi	yuc	United States	35°45N	86°45W	Yuchi	YUC
94. MUSKOGEAN						
Alabama	abm	United States	32°20N	87°25W	Alabama	AKZ
Chickasaw	cck	United States	34°N	88°W	Chickasaw	CIC
Choctaw	cct	United States	32°15N	88°30W	Choctaw	CCT

	Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
	Creek	crk	United States	34°N	85°W	Muskogee	CRK
	Koasati	koa	United States	34°50N	85°10W	Koasati	CKU
	Mikasuki	mki	United States	32°N	85°W	Mikasuki	MIK
	Seminole	smn	United States	33°N	84°W	dialect of Muskogee	dCRK
95. TUNICA[69]							
	Tunica	tun	United States	32°40N	91°W	Tunica	TUK
96. NATCHEZ[70]							
	Natchez	nat	United States	31°45N	91°20W	Natchez	NCZ
97. ATAKAPA[71]							
	Atakapa	atk	United States	30°N	93°30W	Atakapa	ALE
98. CHITIMACHA[72]							
	Chitimacha	ctm	United States	29°40N	91°W	Chitimacha	CHM
99. TONKAWA[73]							
	Tonkawa	ton	United States	30°15N	96°45W	Tonkawa	TON
100. KARANKAWA							
	Karankawa	kkw	United States	28°N	97°W	—	—
101. COAHUILTECAN							
	Coahuilteco	coa	Mexico	28°N	100°W	—	—
102. COMECRUDAN							
	Comecrudo	cmc	Mexico	25°50N	99°W	—	—
103. CADDOAN							
	Arikara	akr	United States	45°30N	100°30W	Arikara	ARI
	Caddo	cad	United States	33°20N	93°30W	Caddo	CAD
	Kitsai	kit	United States	31°50N	96°30W	Kitsai	KII
	Pawnee	pwn	United States	41°N	98°40W	Pawnee	PAW
	Wichita	wic	United States	33°20N	97°20W	Wichita	WIC
104. SIOUAN							
	Biloxi	blx	United States	30°30N	88°40W	Biloxi	BLL
	Catawba	ctw	United States	35°30N	80°30W	Catawba	CHC
	Crow	cro	United States	47°N	108°W	Crow	CRO
	Dakota	dak	United States	45°N	93°30W	Dakota	DHG
	Hidatsa	hid	United States	47°N	102°30W	Hidatsa	HID
	Lakhota	lkt	United States	43°50N	101°50W	Lakota	LKT
	Mandan	mdn	United States	46°30N	102°30W	Mandan	MHQ
	Omaha	omh	United States	42°N	97°15W	dialect of Omaha-Ponca	dOMA
	Osage	osa	United States	37°N	94°W	Osage	OSA
	Oto	oto	United States	40°30N	96°W	dialect of Iowa-Oto	dIOW
	Stoney	sto	Canada	53°45N	116°30W	Stoney	STO
	Tutelo	tte	United States	37°40N	78°45W	Tutelo	TTA
	Winnebago	wnb	United States	43°30N	88°30W	Hocák	WIN

[69] Treated by *Ethnologue* as belonging to Gulf, a family not recognized here.
[70] Treated by *Ethnologue* as belonging to Gulf, a family not recognized here.
[71] Treated by *Ethnologue* as belonging to Gulf, a family not recognized here.
[72] Treated by *Ethnologue* as belonging to Gulf, a family not recognized here.
[73] Treated by *Ethnologue* as belonging to Coahuitecan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
105. KIOWA-TANOAN						
Jemez	jem	United States	35°50N	107°W	Jemez	TOW
Kiowa	kio	United States	37°N	99°W	Kiowa	KIO
Tewa (Arizona)	tew	United States	35°50N	110°25W	Tewa	TEW
Tewa (Rio Grande)	trg	United States	36°15N	106°10W	dialect of Tewa	dTEW
Tewa (San Juan Pueblo)	tsj	United States	36°05N	106°05W	dialect of Tewa	dTEW
Tiwa (Northern)	twn	United States	36°30N	105°30W	Tiwa (Northern)	TAO
Tiwa (Southern)	tws	United States	35°N	106°20W	Tiwa (Southern)	TIX
106. KERESAN						
Acoma	aco	United States	34°55N	107°35W	Keres (Western)	KJQ
Keresan (Santa Ana)	ksa	United States	35°25N	106°40W	dialect of Keres (Eastern)	dKEE
107. ZUNI						
Zuni	zun	United States	35°05N	108°50W	Zuni	ZUN
108. UTO-AZTECAN						
108.1. Aztecan						
Nahuatl (Central)	nhc	Mexico	19°N	98°15W	Nahuatl (Central)	NHN
Nahuatl (Huasteca)	nhh	Mexico	22°N	99°W	2 varieties: Náhuatl (Huasteca Este)/ Náhuatl (Huasteca Oeste)	NAI/NHW
Nahuatl (Huauchinango)	nhu	Mexico	20°10N	98°05W	—	—
Nahuatl (Mecayapan Isthmus)	nmi	Mexico	18°15N	94°50W	Nahuatl (Istmo-Mecayapan)	NAU
Nahuatl (Michoacán)	nhm	Mexico	18°30N	103°W	Náhuatl (Michoacán)	NCL
Nahuatl (Milpa Alta)	nmp	Mexico	19°15N	99°10W	—	—
Nahuatl (North Puebla)	nhn	Mexico	20°N	98°15W	Náhuatl (Puebla Norte)	NCJ
Nahuatl (Pajapan)	npa	Mexico	18°15N	94°45W	Náhuatl (Istmo-Pajapan)	NHP
Nahuatl (Pochutla)	nhp	Mexico	15°45N	96°30W	—	—
Nahuatl (Sierra de Zacapoaxtla)	nsz	Mexico	19°35N	97°20W	Nahuatl (Puebla Sierra)	AZZ
Nahuatl (Tetelcingo)	nht	Mexico	19°40N	99°W	Náhuatl (Tetelcingo)	NHG
Nahuatl (Xalitla)	nhx	Mexico	18°N	99°30W	dialect of Náhuatl (Guerrero)	dNAH
Pipil	pip	El Salvador	13°50N	89°35W	Pipil	PPL
108.2. Cahita						
Eudeve	eud	Mexico	29°10N	109°40W	—	—
Mayo	myo	Mexico	26°N	108°W	Mayo	MAY
Yaqui	yaq	Mexico	27°30N	110°15W	Yaqui	YAQ
108.3. Corachol						
Cora	cor	Mexico	22°10N	104°50W	Cora	COR
Huichol	hui	Mexico	22°N	104°W	Huichol	HCH
108.4. Hopi						
Hopi	hop	United States	36°N	110°W	Hopi	HOP
108.5. Numic						
Chemehuevi	cmh	United States	34°20N	115°10W	dialect of Ute-Southern Paiute	dUTE
Comanche	cmn	United States	33°30N	101°30W	Comanche	COM
Kawaiisu	kws	United States	36°N	117°30W	Kawaiisu	KAW
Mono (in United States)	mno	United States	38°N	119°W	Mono	MON
Paiute (Northern)	pno	United States	42°N	118°W	Paiute (Northern)	PAO
Paiute (Southern)	put	United States	37°50N	113°20W	dialect of Ute-Southern Paiute	dUTE
Shoshone	sho	United States	41°N	114°W	Shoshoni	SHH
Shoshone (Wind River)	swr	United States	43°25N	108°50W	dialect of Shoshoni	dSHH
Tümpisa Shoshone	tsh	United States	37°N	117°W	Panamint	PAR
Ute	ute	United States	39°N	109°W	dialect of Ute-Southern Paiute	dUTE
108.6. Takic						
Cahuilla	cah	United States	33°30N	116°15W	Cahuilla	CHL
Cupeño	cup	United States	33°10N	116°30W	Cupeño	CUP

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Luiseño	lui	United States	33°20N	117°10W	Luiseño	LUI
Serrano	srr	United States	34°30N	117°W	Serrano	SER
108.7. Tarahumaran						
Guarijío	grj	Mexico	27°45N	108°40W	Huarijío	VAR
Tarahumara (Central)	tce	Mexico	27°30N	107°30W	Tarahumara Central	TAR
Tarahumara (Western)	twe	Mexico	27°30N	108°W	Tarahumara Baja	TAC
108.8. Tepiman						
Nevome	nev	Mexico	31°N	112°30W	dialect of O'odham	dPAP
O'odham	ood	Mexico, United States	32°N	112°W	dialect of O'odham	dPAP
Pima Bajo	pba	Mexico	29°N	111°W	2 varieties: Pima Bajo (Chihuahua)/ Pima Bajo (Sonora)	PMB/PIA
Tepecano	tpc	Mexico	21°45N	104°45W	Tepecano	TEP
Tepehuan (Northern)	tpn	Mexico	26°20N	107°W	Tepehuán (Norte)	NTP
Tepehuan (Southeastern)	tps	Mexico	23°N	104°30W	Tepehuán (Sureste)	STP
108.9. Tubar						
Tubar	tub	Mexico	27°N	108°W	Tubar	TBU
108.10. Tübatulabal						
Tübatulabal	tbb	United States	36°N	118°20W	Tübatulabal	TUB
109. WAKASHAN						
109.1. Northern Wakashan						
Haisla	hsl	Canada	54°N	128°45W	Haisla	HAS
Heiltsuk	hei	Canada	52°N	127°30W	Heiltsuk	HEI
Kwakw'ala	kwk	Canada	51°N	127°W	Kwakiutl	KWK
109.2. Southern Wakashan						
Kyuquot	kyq	Canada	50°10N	127°10W	dialect of Nootka	dNOO
Makah	mak	United States	48°20N	124°40W	Makah	MYH
Nitinaht	nit	Canada	48°40N	124°45W	dialect of Nootka	dNOO
Nuuchahnulth	nuu	Canada	49°40N	126°40W	Nootka	NOO
110. SALISHAN						
110.1. Bella Coola						
Bella Coola	bco	Canada	52°30N	126°40W	Bella Coola	BEL
110.2. Central Salish						
Clallam	cla	United States	48°05N	123°45W	Clallam	CLM
Comox	cmx	Canada	50°10N	125°W	Comox	COO
Halkomelem	hlk	Canada	49°30N	122°30W	Halkomelem	HUR
Lummi	lum	United States	48°42N	122°40W	dialect of Salish (Straits)	dSTR
Lushootseed	lus	United States	48°N	122°W	Lushootseed	LUT
Saanich	sch	Canada	48°35N	123°25W	dialect of Salish (Straits)	dSTR
Salish (Samish Straits)	sss	United States	48°35N	122°35W	dialect of Salish (Straits)	dSTR
Salish (Southern Puget Sound)	sps	United States	47°15N	122°30W	Salish (Southern Puget Sound)	SLH
Salish (Straits)	sst	Canada, United States	48°45N	123°15W	Salish (Straits)	STR
Songish	sgs	Canada	48°27N	123°20W	dialect of Salish (Straits)	dSTR
Sooke	soo	Canada	48°20N	123°45W	dialect of Salish (Straits)	dSTR
Squamish	squ	Canada	49°40N	123°10W	Squamish	SQU
Twana	twa	United States	47°40N	122°45W	Twana	TWA
110.3. Interior Salish						
Coeur d'Alene	coe	United States	47°15N	116°30W	Coeur d'Alene	CRD
Columbia-Wenatchi	cwe	United States	47°30N	120°W	Columbia-Wenatchi	COL
Kalispel	kls	United States	48°N	117°W	Kalispel-Pend d'Oreille	FLA
Lillooet	lil	Canada	50°45N	122°W	Lillooet	LIL
Okanagan	oka	Canada, United States	49°30N	118°30W	Okanagan	OKA
Shuswap	shu	Canada	52°N	120°W	Shuswap	SHS
Spokane	spo	United States	47°40N	117°45W	Spokane	SPO
Thompson	tho	Canada	49°45N	121°45W	Thompson	THP

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
110.4. Tillamook						
Tillamook	til	United States	45°20N	123°45W	Tillamook	TIL
110.5. Tsamosan						
Chehalis (Upper)	chl	United States	46°35N	123°W	Chehalis (Upper)	CJH
111. CHIMAKUAN						
Chemakum	cmk	United States	48°05N	122°55W	—	—
Quileute	qui	United States	47°55N	124°15W	Quileute	QUI
112. KUTENAI						
Kutenai	kut	Canada, United States	49°30N	116°W	Kutenai	KUN
113. PENUTIAN[74]						
113.1. Chinookan						
Chinook (Lower)	ckl	United States	46°15N	123°30W	Chinook	dCHH
Chinook (Upper)	cku	United States	45°30N	122°30W	Wasco-Wishram	WAC
113.2. Klamath-Modoc						
Klamath	klm	United States	42°30N	121°30W	Klamath-Modoc	KLA
113.3. Maiduan						
Konkow	knw	United States	39°30N	121°30W	Maidu (Northwest)	MAI
Maidu (Northeast)	mne	United States	40°N	120°40W	Maidu (Northeast)	NMU
Nisenan	nsn	United States	39°N	121°W	Nisenan	NSZ
113.4. Molala						
Molala	mll	United States	44°30N	122°30W	Molala	MBE
113.5. Sahaptian						
Nez Perce	nez	United States	46°N	116°W	Nez Perce	NEZ
Sahaptin (Northern)	shp	United States	46°30N	120°W	2 varieties: Walla Walla/Yakima	WAA/YAK
Sahaptin (Umatilla)	smt	United States	45°40N	118°30W	Umatilla	UMA
113.6. Tsimshianic						
Gitksan	git	Canada	55°20N	127°45W	Gitxsan	GIT
Nisgha	nsg	Canada	55°N	130°W	Nisga'a	NCG
Tsimshian (Coast)	tsi	Canada, United States	52°30N	129°W	dialect of Tsimshian	dTSI
113.7. Utian						
113.7.1. Costanoan						
Costanoan	cos	United States	37°N	122°W	2 varieties: Costanoan (Northern)/ Costanoan (Southern)	CST/CSS
Mutsun	mut	United States	36°50N	121°30W	dialect of Costanoan (Southern)	dCSS
113.7.2. Miwok						
Miwok (Bodega)	mkb	United States	38°20N	123°W	Miwok (Coast)	CSI
Miwok (Central Sierra)	mcs	United States	38°N	120°20W	Miwok (Central Sierra)	CSM
Miwok (Lake)	mwl	United States	38°55N	122°40W	Miwok (Lake)	LMW
Miwok (Northern Sierra)	mwn	United States	38°20N	120°40W	Miwok (Northern Sierra)	NSQ
Miwok (Plains)	mwp	United States	38°N	121°W	Miwok (Plains)	PMW
Miwok (Southern Sierra)	mss	United States	37°30N	120°W	Miwok (Southern Sierra)	SKD

[74] The status of Penutian as a language family, and its membership if it is one, are both controversial. The version of Penutian included here is relatively conservative in excluding a number of groups that *Ethnologue* includes and that are often included but whose membership in Penutian appears to be more controversial than the groups included here. The other groups that *Ethnologue* includes but which are treated here as separate families are Oregon Coast, Takelma, and Kalapuyan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
113.8. Wintuan						
Patwin	ptw	United States	39°N	122°20W	dialect of Wintu	dWIT
Wintu	win	United States	41°N	122°30W	Wintu	WIT
113.9. Yokuts						
Wikchamni	wik	United States	36°25N	119°05W	dialect of Yokuts	dYOK
Yawelmani	ywl	United States	35°25N	119°W	dialect of Yokuts	dYOK
Yokuts (Yaudanchi)	yok	United States	36°05N	119°05W	dialect of Yokuts	dYOK
114. OREGON COAST[75]						
114.1. Alsea						
Alsea	als	United States	44°40N	123°55W	Alsea	AES
114.2. Coosan						
Coos (Hanis)	coo	United States	43°30N	124°10W	Coos	COS
114.3. Siuslawan						
Siuslaw	siu	United States	44°N	124°W	Siuslaw	SIS
115. TAKELMA[76]						
Takelma	tkl	United States	42°30N	123°W	Takelma	TKM
116. KALAPUYAN[77]						
Kalapuya	klp	United States	44°30N	123°W	Kalapuya	KAL
117. HOKAN[78]						
117.1. Chimariko						
Chimariko	chi	United States	41°N	123°W	Chimariko	CID
117.2. Palaihnihan						
Achumawi	acm	United States	41°30N	121°W	Achumawi	ACH
Atsugewi	ats	United States	40°45N	121°W	Atsugewi	ATW
117.3. Pomoan						
Kashaya	ksh	United States	38°40N	123°20W	Kashaya	KJU
Pomo (Central)	pmc	United States	39°N	123°20W	Pomo (Central)	POO
Pomo (Eastern)	pme	United States	39°N	122°40W	Pomo (Eastern)	PEB
Pomo (Northern)	pmn	United States	39°20N	123°30W	Pomo (Northern)	PEJ
Pomo (Southeastern)	pso	United States	39°N	122°30W	Pomo (Southeastern)	PEO
117.4. Seri						
Seri	ser	Mexico	29°N	112°W	Seri	SEI
117.5. Shasta						
Shasta	shs	United States	41°50N	122°40W	Shasta	SHT
117.6. Yana						
Yana	yan	United States	40°30N	122°W	Yana	YNN

[75] The status of Oregon Coast as a family is controversial. On a conservative classification, the three genera included would each be separate families. In *Ethnologue*, these languages are treated as belonging to Penutian.

[76] Treated by *Ethnologue* as belonging to Penutian.

[77] Treated by *Ethnologue* as belonging to Penutian.

[78] The status of Hokan as a language family, and its membership if it is one, are both controversial. The version of Hokan included here is relatively conservative in excluding a number of groups that *Ethnologue* includes and that are often included but whose membership in Hokan appears to be more controversial than the groups included here. The other groups that *Ethnologue* includes but which are treated here as separate families are Esselen, Karok, Salinan, Tequistlatecan, and Washo.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
117.7. Yuman						
Cocopa	ccp	United States, Mexico	32°20N	115°W	Cocopa	COC
Diegueño (Mesa Grande)	die	Mexico, United States	32°40N	116°10W	dialect of Kumiái	dDIH
Havasupai	hav	United States	35°45N	112°30W	dialect of Havasupai-Walapai-Yavapai	dYUF
Hualapai	hlp	United States	35°30N	113°45W	dialect of Havasupai-Walapai-Yavapai	dYUF
Kiliwa	klw	Mexico	31°20N	115°40W	Kiliwa	KLB
Maricopa	mar	United States	33°10N	113°10W	Maricopa	MRC
Mojave	moj	United States	34°40N	114°35W	Mohave	MOV
Quechan	que	United States	32°50N	114°20W	Quechan	YUM
Tiipay (Jamul)	tja	United States, Mexico	32°10N	116°30W	dialect of Kumiái	dDIH
Yavapai	yav	United States	34°N	113°20W	dialect of Havasupai-Walapai-Yavapai	dYUF
118. KAROK[79]						
Karok	krk	United States	41°40N	123°W	Karok	KYH
119. WASHO[80]						
Washo	was	United States	39°15N	120°W	Washo	WAS
120. WAPPO-YUKIAN						
120.1. Wappo						
Wappo	wap	United States	38°30N	122°30W	Wappo	WAO
120.2. Yukian						
Yuki	yki	United States	39°40N	123°30W	Yuki	YUK
121. SALINAN[81]						
Salinan	sal	United States	36°N	121°W	Salinan	SAL
122. ESSELEN[82]						
Esselen	ess	United States	36°15N	121°45W	Esselen	ESQ
123. CHUMASH						
Chumash (Barbareño)	cba	United States	34°30N	120°15W	Barbareño	BOI
Chumash (Ineseño)	cin	United States	34°40N	120°15W	Ineseño	INE
124. OTO-MANGUEAN						
124.1. Amuzgoan						
Amuzgo	amz	Mexico	16°50N	98°W	3 varieties: Amuzgo (Guerrero)/. . .	AMU
124.2. Chichimec						
Chichimeca-Jonaz	cjo	Mexico	21°40N	100°30W	Chichimeca-Jonaz	PEI
124.3. Chinantecan						
Chinantec (Comaltepec)	ccm	Mexico	17°35N	96°25W	Chinanteco (Comaltepec)	CCO
Chinantec (Lealao)	cle	Mexico	17°20N	95°55W	Chinanteco (Lealao)	CLE
Chinantec (Palantla)	cpl	Mexico	18°50N	96°45W	Chinanteco (Palantla)	CPA
Chinantec (Quiotepec)	chq	Mexico	17°35N	96°40W	Chinanteco (Quiotepec)	CHQ
Chinantec (San Felipe Usila)	csf	Mexico	17°55N	96°30W	Chinanteco (Usila)	CUS
Chinantec (Sochiapan)	csc	Mexico	17°45N	96°40W	Chinanteco (Sochiapan)	CSO
Chinantec (Tepetotutla)	cte	Mexico	17°50N	96°30W	Chinanteco (Tepetotutla)	CNT

[79] Treated by *Ethnologue* as belonging to Hokan.
[80] Treated by *Ethnologue* as belonging to Hokan.
[81] Treated by *Ethnologue* as belonging to Hokan.
[82] Treated by *Ethnologue* as belonging to Hokan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
124.4. Mixtecan						
Cuicatec	ctc	Mexico	17°50N	96°50W	2 varieties: Cuicateco (Tepeuxila)/ Cuicateco (Teutila)	CUX/CUT
Mixtec (Alacatlatzala)	mxl	Mexico	17°15N	98°35W	Mixteco (Alacatlatzala)	MIM
Mixtec (Atatlahuca)	mxa	Mexico	17°N	97°45W	Mixteco (Atatláhuca)	MIB
Mixtec (Ayutla)	mxt	Mexico	16°55N	99°10W	Mixteco (Ayutla)	MIY
Mixtec (Chalcatongo)	mxc	Mexico	17°03N	97°35W	dialect of Mixteco (San Miguel el Grande)	dMIG
Mixtec (Chayuco)	mxu	Mexico	16°25N	97°50W	Mixteco (Chayuco)	MIH
Mixtec (Coatzospan)	mxz	Mexico	18°05N	96°35W	Mixteco (Coatzospan)	MIZ
Mixtec (Jamiltepec)	mja	Mexico	16°15N	97°50W	Mixteco (Jamiltepec)	MXT
Mixtec (Jicaltepec)	mxj	Mexico	16°20N	98°W	Mixteco (Pinotepa Nacional)	MIO
Mixtec (Molinos)	mxm	Mexico	17°N	97°35W	dialect of Mixteco (San Miguel el Grande)	dMIG
Mixtec (Ocotepec)	mxo	Mexico	17°10N	97°45W	Mixteco (Ocotepec)	MIE
Mixtec (Peñoles)	mxp	Mexico	17°05N	96°55W	Mixteco (Peñoles)	MIL
Mixtec (San Juan Colorado)	mjc	Mexico	16°30N	98°W	Mixteco (San Juan Colorado)	MJC
Mixtec (San Miguel el Grande)	mxg	Mexico	17°03N	97°34W	Mixteco (San Miguel el Grande)	MIG
Mixtec (Silacayoapan)	mxs	Mexico	17°30N	98°10W	Mixteco (Silacayoapan)	MKS
Mixtec (Yosondúa)	mxy	Mexico	16°55N	97°35W	Mixteco (Yosondúa)	MPM
Trique (Chicahuaxtla)	trc	Mexico	17°10N	97°50W	Trique (Chicahuaxtla)	TRS
Trique (Copala)	tri	Mexico	17°10N	97°55W	Trique (Copala)	TRC
124.5. Otomian						
Mazahua	maz	Mexico	19°25N	99°55W	2 varieties: Mazahua (Central)/ Mazahua (Michoacán)	MAZ/QMN
Ocuilteco	ocu	Mexico	18°55N	99°40W	Matlatzinca (Atzingo)	OCU
Otomí (Ixtenco)	oix	Mexico	19°05N	97°55W	Otomi (Ixtenco)	OTA
Otomí (Mezquital)	otm	Mexico	20°10N	99°10W	Otomí (Mezquital)	OTE
Otomí (Santiago Mexquititlan)	osm	Mexico	20°05N	100°05W	dialect of Otomi (Northwestern)	dOTQ
Otomí (Sierra)	ots	Mexico	20°40N	98°50W	Otomi (Sierra Oriental)	OTM
124.6. Pamean						
Pame	pam	Mexico	22°N	99°30W	3 varieties: Pame (Central)/. . .	PBS/. . .
124.7. Popolocan						
Chocho	cch	Mexico	17°40N	97°25W	Chochoteco	COZ
Ixcatec	ixc	Mexico	17°50N	97°10W	Ixcateco	IXC
Mazatec (Chiquihuitlán)	mzc	Mexico	17°45N	96°55W	Mazatec (Chiquihuitlán)	MAQ
Mazatec (Huautla)	mzh	Mexico	18°15N	96°50W	Mazateco (Huautla)	MAU
Popoloca (Metzontla)	pop	Mexico	18°10N	97°30W	Popoloca (Mezontla)	PBE
Popoloca (San Juan Atzingo)	psj	Mexico	18°10N	96°25W	Popoloca (San Juan Atzingo)	POE
Popoloca (San Vicente Coyotepec)	psv	Mexico	18°15N	97°50W	dialect of Popoloca (Coyotepec)	dPBF
124.8. Zapotecan						
Chatino (Nopala)	cht	Mexico	16°05N	97°10W	Chatino (Nopala)	CYA
Chatino (Sierra Occidental)	cso	Mexico	16°15N	97°20W	Chatino (Sierra Occidental)	CTP
Chatino (Tataltepec)	ctt	Mexico	16°10N	97°35W	Chatino (Tataltepec)	CTA
Chatino (Yaitepec)	cya	Mexico	16°15N	97°15W	dialect of Chatino (Sierra Occidental)	dCTP
Zapotec (Isthmus)	zai	Mexico	16°20N	94°10W	Zapoteco (Istmo)	ZAI
Zapotec (Ixtlan)	zpi	Mexico	17°15N	96°30W	Zapateco (Ixtlán Sureste)	ZPD
Zapotec (Juárez)	zaj	Mexico	17°30N	96°35W	Zapoteco (Juárez Sierra)	ZAA
Zapotec (Mitla)	zap	Mexico	16°48N	96°15W	Zapoteco (Mitla)	ZAW
Zapotec (Mixtepec)	zam	Mexico	16°20N	96°20W	Zapoteco (Mixtepec)	ZPM
Zapotec (Quiegolani)	zaq	Mexico	16°10N	96°05W	Zapoteco (Santa María Quiegolani)	ZPI
Zapotec (San Lucas Quiaviní)	zsq	Mexico	16°54N	96°28W	dialect of Zapotec (San Juan Guelavía)	dZAB
Zapotec (Texmelucan)	zte	Mexico	16°30N	97°10W	Zapoteco (Texmelucan)	ZPZ
Zapotec (Yatzachi)	zya	Mexico	17°12N	96°12W	Zapotec (Yatzachi)	ZAV
Zapotec (Zoogocho)	zzo	Mexico	17°15N	96°15W	Zapotec (Zoogocho)	ZPQ
125. TARASCAN						
Purépecha	pur	Mexico	19°30N	101°40W	Purépecha	TSZ

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
126. *TOTONACAN*						
Tepehua	tep	Mexico	20°50N	98°10W	3 varieties: Tepehua (Huehuetla)/. . .	TEE/. . .
Totonac (Misantla)	tot	Mexico	19°55N	96°55W	Totonaca (Yecuatla)	TLC
Totonac (Papantla)	tpa	Mexico	20°20N	97°20W	Totonaca (Papantla)	TOP
Totonac (Sierra)	tos	Mexico	19°55N	97°25W	Totonac (Sierra)	TOS
Totonac (Xicotepec de Juárez)	txj	Mexico	20°25N	97°50W	Totonaca (Xicotepec de Juarez)	TOO
127. *CUITLATEC*						
Cuitlatec	cut	Mexico	17°30N	101°W	—	—
128. *SUBTIABA-TLAPANEC*						
Tlapanec	tlp	Mexico	17°05N	99°W	3 varieties: Tlapaneco (Acatepec)/. . .	TPX/. . .
129. *TEQUISTLATECAN*[83]						
Chontal (Highland)	cho	Mexico	16°15N	95°45W	Chontal de Oaxaca (Sierra)	CHD
Chontal (Huamelultec Oaxaca)	chx	Mexico	16°N	95°45W	Chontal de Oaxaca (Costa)	CLO
130. *MIXE-ZOQUE*						
Mixe (Coatlán)	mco	Mexico	16°55N	95°40W	Mixe (Coatlán)	MCO
Mixe (Tlahuitoltepec)	mth	Mexico	17°05N	96°10W	Mixe (Tlahuitoltepec)	MXP
Mixe (Totontepec)	mtp	Mexico	17°15N	96°W	Mixe (Totontepec)	MTO
Olutec	olu	Mexico	17°57N	95°W	Popoluca (Oluta)	PLO
Sayultec	say	Mexico	17°50N	95°W	Popoluca (Sayula)	POS
Zoque (Chimalapa)	zch	Mexico	16°45N	94°45W	Zoque (Chimalapa)	ZOH
Zoque (Copainalá)	zqc	Mexico	17°N	93°15W	Zoque (Copainalá)	ZOC
Zoque (Francisco León)	zfl	Mexico	17°20N	93°15W	Zoque (Francisco León)	ZOS
Zoque (Ostuacan)	zqo	Mexico	17°25N	93°18W	dialect of Zoque (Copainalá)	dZOC
Zoque (Rayon)	zqr	Mexico	17°05N	93°W	Zoque (Rayon)	ZOR
Zoque (Soteapan)	zqs	Mexico	18°20N	95°10W	Popoluca (Sierra)	POI
131. *HUAVEAN*						
Huave (San Mateo del Mar)	hve	Mexico	16°13N	95°W	Huave (San Mateo del Mar)	HUV
132. *MAYAN*						
Achí	aci	Guatemala	15°10N	90°30W	2 varieties: Achí (Cubulco)/ Achí (Rabinal)	ACC/ACR
Aguacatec	agu	Guatemala	15°25N	91°20W	Aguacateco	AGU
Cakchiquel	cak	Guatemala	14°30N	91°W	10 varieties: Cakchiquel (Southern)/. . .	CKF/. . .
Chicomuceltec	cec	Mexico	15°30N	92°15W	Chicomuceltec	COB
Chol	col	Mexico	17°45N	92°30W	2 varieties: Ch'ol (Tila)/Ch'ol (Tumbalá)	CTI/CTU
Chontal Maya	cmy	Mexico	18°10N	92°35W	Chontal (Tabasco)	CHF
Chortí	coi	Guatemala, Honduras	14°50N	89°15W	Chortí	CAA
Chuj	chj	Guatemala	15°55N	91°35W	2 varieties: Chuj (San Mateo Ixtatán)/ Chuj (San Sebastián Coatán)	CNM/CAC
Huastec	htc	Mexico	22°05N	99°20W	3 varieties: Huasteco (San Francisco Chontla)/. . .	HAU/. . .
Itzaj	itz	Guatemala	17°N	89°50W	Itzá	ITZ
Ixil	ixi	Guatemala	15°30N	91°W	2 varieties: Ixil (Nebaj)/Ixil (San Juan Cotzal)	IXI/IXL
Jakaltek	jak	Guatemala	15°40N	91°40W	2 varieties: Jacalteco (Eastern)/ Jacalteco (Western)	JAC/JAI
Kanjobal (Eastern)	kea	Guatemala	15°20N	91°40W	Kanjobal (Eastern)	KJB
Kanjobal (Western)	kwe	Guatemala	15°50N	91°50W	Kanjobal (Western)	KNJ
K'ekchí	kek	Guatemala	16°N	89°50W	Kekchí	KEK

[83] Treated by *Ethnologue* as belonging to Hokan.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Lacandón	lac	Mexico	17°N	91°30W	Lacandón	LAC
Mam	mam	Guatemala	15°N	91°50W	5 varieties: Mam (Central)/. . .	MVC/. . .
Mam (Southern)	mmz	Guatemala	14°50N	91°40W	Mam (Southern)	MMS
Mopan	mop	Belize	16°35N	88°40W	Mopán Maya	MOP
Pocomam	pcm	Guatemala	14°40N	90°30W	3 varieties: Pokomam (Central)/. . .	POC/. . .
Pocomam Oriental	poc	Guatemala	14°45N	89°45W	Pokomam (Eastern)	POA
Pokomchí	pkm	Guatemala	15°25N	90°30W	2 varieties: Pokomchí (Eastern)/ Pokomchí (Western)	POH/POB
Quiché	qch	Guatemala	15°N	91°25W	6 varieties: Quiché (Central)/. . .	QUC/. . .
Teco	tec	Guatemala, Mexico	15°25N	92°05W	Tectiteco	TTC
Tojolabal	toj	Mexico	16°20N	91°30W	Tojolabal	TOJ
Tzeltal	tze	Mexico	16°50N	92°15W	2 varieties: Tzeltal (Bachajón)/ Tzeltal (Oxchuc)	TZB/TZH
Tzeltal (Aguacatenango)	tza	Mexico	16°25N	92°30W	dialect of Tzeltal (Oxchuc)	dTZH
Tzeltal (Bachajón)	tzb	Mexico	16°55N	92°W	Tzeltal (Bachajón)	TZB
Tzeltal (Tenejapa)	tzt	Mexico	16°40N	92°20W	dialect of Tzeltal (Oxchuc)	dTZH
Tzotzil	tzo	Mexico	16°30N	92°40W	6 varieties: Tzotzil (Chamula)/. . .	TZC/. . .
Tzotzil (San Andrés)	tzs	Mexico	17°N	92°50W	Tzotzil (San Andrés Larrainzar)	TZS
Tzotzil (Zinacantán)	tzz	Mexico	16°50N	92°50W	Tzotzil (Zinacantán)	TZZ
Tzutujil	tzu	Guatemala	14°40N	91°20W	2 varieties: Tzutujil (Eastern)/ Tzutujil (Western)	TZJ/TZT
Yucatec	yct	Mexico	20°N	89°W	Maya (Yucatán)	YUA
133. TOL						
Tol	tol	Honduras	14°40N	87°W	Tol	JIC
134. MISUMALPAN						
Miskito	mis	Nicaragua	14°N	83°40W	Mískito	MIQ
Sumu	sum	Nicaragua	13°N	84°30W	Sumo Tawahka	SUM
135. CHIBCHAN						
135.1. Aruak						
Cogui	cog	Colombia	11°N	73°50W	Cogui	KOG
Damana	dam	Colombia	11°N	73°30W	Malayo	MBP
Ika	ika	Colombia	10°40N	73°45W	Ica	ARH
135.2. Chibchan Proper						
Muisca	msc	Colombia	5°N	74°W	Chibcha	CBF
Tunebo	tnb	Colombia	6°45N	72°15W	4 varieties: Tunebo (Angosturas)/. . .	TND/. . .
135.3. Guaymi						
Buglere	bgl	Panama	8°30N	81°15W	Buglere	SAB
Ngäbere	nbr	Panama	8°40N	82°W	Ngäbere	GYM
135.4. Kuna						
Kuna	kun	Panama, Colombia	8°N	77°20W	Kuna (Border)	KUA
135.5. Motilon						
Motilón (Chibchan)	mti	Colombia, Venezuela	9°N	73°W	Motilón	MOT
135.6. Paya						
Pech	pec	Honduras	15°N	85°30W	Pech	PAY
135.7. Rama						
Rama	ram	Nicaragua	11°45N	83°45W	Rama	RMA
135.8. Talamanca						
Boruca	brc	Costa Rica	8°50N	83°15W	Boruca	BRN
Bribri	bri	Costa Rica	9°25N	83°W	Bribri	BZD
Cabécar	cab	Costa Rica	9°45N	83°25W	Cabécar	CJP
Teribe	trb	Costa Rica, Panama	9°20N	82°40W	Teribe	TFR

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
136. CHOCO						
Catio	cat	Colombia	7°10N	76°20W	Emberá-Catío	CTO
Embera	emb	Panama, Colombia	6°50N	77°10W	Emberá (Northern)	EMP
Embera Chami	emc	Colombia	5°N	76°W	Emberá-Chamí	CMI
Epena Pedee	epe	Colombia	3°N	77°W	Emberá-Saija	SJA
Waunana	wau	Colombia, Panama	4°N	77°W	Woun Meu	NOA
137. CHIMILA						
Chimila	cma	Colombia	10°N	74°W	Chimila	CBG
138. YURIMANGÍ						
Yurimangí	yrm	Colombia	3°50N	77°W	—	—
139. PÁEZAN						
Páez	pae	Colombia	2°40N	76°W	Páez	PBB
140. BARBACOAN						
Awa Pit	awp	Colombia, Ecuador	1°30N	78°15W	Awa-Quaiquer	KWI
Cayapa	cay	Ecuador	0°40N	79°W	Chachi	CBI
Guambiano	gmb	Colombia	2°30N	76°40W	Guambiano	GUM
Tsafiki	tsf	Ecuador	1°S	79°20W	Colorado	COF
141. CAMSÁ						
Camsá	cam	Colombia	1°10N	77°W	Camsá	KBH
142. GUAHIBAN						
Cuiba	cui	Colombia	6°30N	70°W	Cuiba	CUI
Guahibo	ghb	Colombia	5°N	69°W	Guahibo	GUH
Guayabero	gyb	Colombia	2°50N	72°W	Guayabero	GUO
Macaguán	mcg	Colombia	6°30N	71°20W	Macaguán	MBN
Playero	pla	Colombia	7°N	71°W	Playero	GOB
Sikuani	skn	Colombia	6°15N	71°30W	dialect of Guahibo	dGUH
143. SÁLIBAN[84]						
Piaroa	pia	Venezuela	5°N	67°W	Piaroa	PID
Sáliba (in Colombia)	sba	Colombia, Venezuela	6°N	70°W	Sáliba	SLC
144. PUINAVE						
Puinave	pui	Colombia, Venezuela	4°N	68°W	Puinave	PUI
145. TUCANOAN						
Barasano	brs	Colombia	0°10S	70°40W	Barasana	BSN
Barasano (Northern)	bno	Colombia	0°20N	70°15W	Waimaha	BAO
Carapana	crp	Colombia, Brazil	0°50N	70°45W	Carapana	CBC
Cubeo	cub	Colombia	1°20N	70°30W	Cubeo	CUB
Desano	des	Brazil, Colombia	0°50N	69°50W	Desano	DES
Guanano	gno	Brazil, Colombia	0°50N	69°30W	Guanano	GVC
Koreguaje	kje	Colombia	1°N	75°30W	Koreguaje	COE
Macuna	mcn	Colombia, Brazil	0°20S	70°10W	Macuna	MYY
Orejón	ore	Peru	2°50S	72°30W	Orejón	ORE
Piratapuyo	prt	Brazil, Colombia	0°20N	69°30W	Piratapuyo	PIR
Retuarã	ret	Colombia	0°30S	70°40W	Tanimuca-Retuarã	TNC

[84] Called Salivan by *Ethnologue*.

	Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
	Secoya	sec	Ecuador, Peru	0°20S	76°W	Secoya	SEY
	Siona	sin	Colombia, Ecuador	0°20N	76°15W	Siona	SIN
	Siriano	sri	Colombia, Brazil	0°45N	70°10W	Siriano	SRI
	Tatuyo	tty	Colombia	0°25N	70°30W	Tatuyo	TAV
	Tucano	tuc	Brazil, Colombia	0°30N	69°10W	Tucano	TUO
	Tuyuca	tuy	Colombia, Brazil	0°30N	70°05W	Tuyuca	TUE
	Yuruti	yrt	Colombia	1°N	70°25W	Yuruti	YUI
146. CACUA-NUKAK[85]							
	Cacua	cac	Colombia	1°05N	70°W	Cacua	CBV
	Nukak	nuk	Colombia	2°30N	71°30W	Nukak Makú	MBR
147. ANDOKE							
	Andoke	adk	Colombia	0°40S	72°W	Andoque	ANO
148. HUITOTOAN							
148.1. Boran							
	Muinane	mui	Colombia	1°S	72°30W	Muinane	BMR
148.2. Huitoto							
	Bora	bor	Peru, Colombia	2°10S	72°20W	Bora	BOA
	Huitoto	htt	Colombia, Peru	1°S	74°W	3 varieties: Huitoto (Minica)/. . .	HTO/. . .
	Huitoto (Minica)	hmi	Colombia	0°20S	74°W	Huitoto (Minica)	HTO
	Huitoto (Muinane)	hmu	Peru	2°25S	71°W	Huitoto (Nipode)	HUX
	Huitoto (Murui)	hum	Colombia, Peru, Brazil	1°S	73°30W	Huitoto (Murui)	HUU
	Ocaina	oca	Peru	2°45S	71°45W	Ocaina	OCA
149. JIRAJARAN							
	Ayomán	ayo	Venezuela	11°N	69°W	—	—
150. TIMOTE-CUICA							
	Cuica	cuc	Venezuela	9°N	70°30W	—	—
151. BETOI							
	Betoi	bti	Venezuela, Colombia	7°10N	71°15W	—	—
152. YARURO							
	Yaruro	yrr	Venezuela	7°N	68°W	Yaruru	YAE
153. WARAO							
	Warao	wra	Venezuela	9°20N	61°40W	Warao	WBA
154. YANOMAM							
	Sanuma	snm	Venezuela, Brazil	4°30N	64°40W	Sanumá	SAM
	Shiriana	shi	Brazil, Venezuela	3°30N	62°50W	Ninam	SHB
	Yanomámi	ynm	Brazil	2°20N	63°W	Yanomámi	WCA
155. COFÁN							
	Cofán	cof	Ecuador, Colombia	0°10N	77°10W	Cofán	CON

[85] *Ethnologue* treats these languages, as well as those in Vaupés-Japurá, as belonging to a larger family, Maku.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
156. TACAME						
Esmeraldeño	esm	Ecuador	0°20N	79°50W	—	—
157. WAORANI						
Waorani	wao	Ecuador	1°S	76°30W	Waorani	AUC
158. ZAPAROAN						
Arabela	arb	Peru	2°S	75°10W	Arabela	ARL
Iquito	iqu	Peru	3°15S	74°W	Iquito	IQU
Zaparo	zpr	Ecuador, Peru	2°S	76°20W	Záparo	ZRO
159. PEBA-YAGUAN						
Yagua	yag	Peru	3°30S	72°W	Yagua	YAD
160. TICUNA						
Ticuna	tic	Brazil, Peru	4°S	70°30W	Ticuna	TCA
161. TAUSHIRO						
Taushiro	tsr	Peru	2°30S	76°W	Taushiro	TRR
162. JIVAROAN						
Achuar	acu	Peru	3°S	76°W	Achuar-Shiwiar	ACU
Aguaruna	agr	Peru	5°S	78°W	Aguaruna	AGR
Huambisa	hmb	Peru	4°S	78°W	Huambisa	HUB
Jivaro	jiv	Ecuador	2°30S	78°W	Shuar	JIV
163. CANDOSHI						
Candoshi	can	Peru	4°10S	77°W	Candoshi-Shapra	CBU
164. URARINA						
Urarina	urn	Peru	4°S	75°W	Urarina	URA
165. CAHUAPANAN						
Chayahuita	chy	Peru	5°30S	77°W	Chayahuita	CBT
Jebero	jeb	Peru	5°25S	76°30W	Jebero	JEB
166. CHIMÚAN						
Mochica	mcc	Peru	7°30S	79°20W	—	—
167. CHOLON						
Cholon	cln	Peru	8°S	77°30W	Cholon	CHT
168. PANOAN						
Amahuaca	amc	Peru, Brazil	10°30S	72°30W	Amahuaca	AMC
Capanahua	cap	Peru	7°S	74°W	Capanahua	KAQ
Cashibo	cas	Peru	8°30S	75°30W	dialect of Cashibo-Cacataibo	dCBR
Cashinahua	csh	Peru, Brazil	10°30S	71°50W	Cashinahua	CBS
Chácobo	cbo	Bolivia	12°10S	66°45W	Chacobo	CAO
Karipuna do Guapore	kdg	Brazil	10°S	65°W	dialect of Karipuná	dKUQ

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Matis	mts	Brazil	4°25S	70°15W	Matís	MPQ
Matsés	myr	Peru, Brazil	5°25S	73°15W	Matsés	MCF
Sharanahua	shh	Peru, Brazil	11°S	70°W	Sharanahua	MCD
Shipibo-Konibo	shk	Peru	7°30S	75°W	Shipibo-Conibo	SHP
Yaminahua	yam	Peru, Brazil	8°S	73°W	Yaminahua	YAA
169. QUECHUAN						
Inga	iga	Colombia	1°N	77°W	Inga	INB
Quechua (Ancash)	qan	Peru	8°30S	77°30W	6 varieties: Quechua (Ancash Chiquian)/. . .	QEC/. . .
Quechua (Ayacucho)	qay	Peru	14°S	74°W	Quechua (Ayacucho)	QUY
Quechua (Bolivian)	qbo	Bolivia	20°S	66°W	2 varieties: Quechua (North Bolivian)/Quechua (South Bolivian)	QUL/QUH
Quechua (Cajamarca)	qca	Peru	7°S	78°30W	Quechua (Cajamarca)	QNT
Quechua (Cochabamba)	qco	Bolivia	17°30S	66°W	dialect of Quechua (South Bolivian)	dQUH
Quechua (Cuzco)	qcu	Peru	14°30S	71°W	Quechua (Cuzco)	QUZ
Quechua (Ecuadorean)	qec	Ecuador	1°S	79°W	9 varieties: Quichua (Highland, Chimborazo)/. . .	QUG/. . .
Quechua (Huallaga)	qhu	Peru	6°S	76°W	Quechua (Huallaga Huánuco)	QEJ
Quechua (Imbabura)	qim	Ecuador	0°20N	78°W	Quichua (Highland, Imbabura)	QHO
Quechua (Tarma)	qta	Peru	11°25S	75°45W	Quechua (Junín)	QJU
170. AYMARAN						
Aymara	aym	Bolivia, Peru, Chile	17°S	69°W	2 varieties: Aymara (Central)/ Aymara (Southern)	AYM/AYC
Jaqaru	jaq	Peru	13°S	76°W	Jaqaru	JQR
171. HARAKMBET						
Amarakaeri	amk	Peru	12°30S	70°30W	Amarakaeri	AMR
172. MÁKU[86]						
Máku	mkw	Brazil	0°25N	69°50W	—	—
173. VAUPÉS-JAPURÁ[87]						
Dâw	daw	Brazil	0°15S	67°05W	Kamã	KWA
Hupda	hpd	Brazil, Colombia	0°10N	69°15W	Hupdë	JUP
Nadëb	nad	Brazil	1°S	66°30W	Nadëb	MBJ
174. ARAWAKAN						
Achagua	acg	Colombia	4°25N	72°15W	Achagua	ACA
Aikaná	aik	Brazil	12°40S	60°40W	Tubarão	TBA
Amuesha	amu	Peru	10°30S	75°25W	Yanesha'	AME
Apurinã	apu	Brazil	9°S	67°W	Apurinã	APU
Arawak	ara	Suriname	5°30N	55°10W	Arawak	ARW
Bahuana	bah	Brazil	0°30S	62°30W	—	—
Baniwa	bnw	Brazil, Venezuela	2°40N	66°50W	Baniwa	BAI
Baré	bae	Venezuela, Brazil	1°N	67°W	Baré	BAE
Baure	baq	Bolivia	13°05S	64°10W	Baure	BRG
Campa (Axininca)	cax	Peru	12°S	74°W	Asháninca	CNI
Campa Pajonal Asheninca	cpa	Peru	10°40S	74°15W	Ashéninca Pajonal	CJO
Curripaco	cur	Colombia	2°30N	68°30W	Curripaco	KPC
Garífuna	grf	Honduras, Guatemala, Belize	15°40N	88°W	Garífuna	CAB
Goajiro	goa	Colombia, Venezuela	12°N	72°W	Wayuu	GUC
Ignaciano	ign	Bolivia	15°10S	65°25W	Ignaciano	IGN
Iranxe	irx	Brazil	13°S	58°W	Irántxe	IRA
Machiguenga	mch	Peru	12°S	72°40W	Machiguenga	MCB
Maipure	mpr	Venezuela	5°30N	67°30W	—	—
Nomatsiguenga	nom	Peru	11°40S	74°30W	Nomatsiguenga	NOT
Palikur	plk	Brazil	3°N	51°W	Palikúr	PAL
Parecis	prc	Brazil	14°S	57°W	Parecís	PAB
Piapoco	ppc	Colombia	4°N	69°30W	Piapoco	PIO

[86] This is a language isolate apparently not included in *Ethnologue*. It should not be confused with what *Ethnologue* calls Maku (and which can be spelled "Makú", with an acute accent on the last syllable). *Ethnologue*'s Maku is treated here as two separate families, Vaupés-Japurá and Cacua-Nukak.

[87] *Ethnologue* treats these languages, as well as those in Cacua-Nukak, as belonging to a larger family Maku.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Piro	pir	Peru	11°S	73°W	Yine	PIB
Resígaro	res	Colombia, Peru	2°25S	71°30W	Resígaro	RGR
Tariana	tar	Brazil	1°N	69°10W	Tariano	TAE
Terêna	trn	Brazil	20°S	56°W	Terêna	TEA
Wapishana	wps	Brazil, Guyana	2°40N	60°W	Aruma	WAP
Warekena	wrk	Venezuela, Brazil, Colombia	1°30N	67°30W	Guarequena	GAE
Waurá	wur	Brazil	12°30S	53°30W	Waurá	WAU
Yucuna	ycn	Colombia	0°45S	71°W	Yucuna	YCN

175. CARIBAN

Akawaio	akw	Guyana	6°N	59°30W	Akawaio	ARB
Apalaí	apl	Brazil	0°S	54°W	Apalaí	APA
Bakairí	bki	Brazil	14°S	55°W	Bakairí	BKQ
Carib	car	Suriname	5°30 N	56°W	Carib	CRB
Carib (De'kwana)	cde	Venezuela	5°30N	65°W	Maquiritari	MCH
Carijona	crj	Colombia	1°N	73°W	Carijona	CBD
Guaque	guq	Colombia	1°N	72°05W	—	—
Hianacoto	hia	Colombia	0°50N	71°W	dialect of Carijona	dCBD
Hixkaryana	hix	Brazil	1°S	59°W	Hixkaryána	HIX
Japreria	jpr	Venezuela	10°30N	73°W	Japrería	JRU
Macushi	mac	Guyana, Brazil, Venezuela	4°N	60°W	Macushi	MBC
Mapoyo	mpy	Venezuela	6°N	67°W	Mapoyo	MCG
Panare	pnr	Venezuela	6°30N	66°W	Panare	PBH
Pemon	pem	Venezuela, Guyana	5°20N	62°W	Pemon	AOC
Tiriyo	tir	Suriname, Brazil	3°15N	55°45W	Trió	TRI
Umaua	umu	Colombia	1°N	72°W	dialect of Carijona	dCBD
Wai Wai	wai	Brazil, Guyana	1°N	59°W	Waiwai	WAW
Wayana	wyn	Brazil, Suriname, French Guiana	3°15N	54°10W	Wayana	WAY
Yukpa	ykp	Colombia, Venezuela	10°10N	72°45W	Yukpa	YUP

176. TUPIAN

176.1. Arikem

Karitiâna	kna	Brazil	9°30S	64°W	Karitiâna	KTN

176.2. Monde

Gavião	gav	Brazil	10°40S	62°W	Gavião do Jiparaná	GVO

176.3. Munduruku

Mundurukú	muu	Brazil	7°S	58°W	Mundurukú	MYU

176.4 Ramarama

Karó (Arára)	kaa	Brazil	10°20S	62°W	Arára (Rondônia)	ARR

176.5. Tupari

Mekens	mek	Brazil	12°30S	61°30W	Mekem	XME

176.6. Tupi-Guaraní

Aché	ach	Paraguay	25°15S	55°10W	Aché	GUQ
Asuriní	asu	Brazil	3°30S	49°30W	Asuriní	ASU
Chiriguano	crg	Argentina, Paraguay, Bolivia	23°40S	64°20W	Guaraní (Western Argentine)	GUI
Cocama	coc	Peru	5°S	74°30W	Cocama-Cocamilla	COD
Guajajara	gjj	Brazil	5°S	46°W	Guajajára	GUB
Guaraní	gua	Paraguay	26°S	56°W	Guaraní (Paraguayan)	GUG
Júma	jum	Brazil	7°30S	64°W	Júma	JUA
Juruna	jrn	Brazil	5°S	54°30W	Jurúna	JUR
Kaiwá	kaw	Brazil, Paraguay	25°S	55°W	Kaiwá	KGK
Kamaiurá	kma	Brazil	12°S	52°40W	Kamayurá	KAY
Omagua	oma	Peru	4°S	73°30W	Omagua	OMG
Sirionó	srn	Bolivia	15°35S	64°W	Sirionó	SRQ
Tapieté	tpt	Paraguay	19°45S	61°40W	Tapieté	TAI
Tupi	tup	Brazil	6°S	36°W	—	—
Urubú-Kaapor	urk	Brazil	2°20S	46°30W	Urubú-Kaapor	URB
Wayampi	way	Brazil	1°N	52°W	Wayampi (Amapari)/Wayampi (Oiapoque)	OYA/ OYM

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
177. MACRO-GE						
177.1. Bororo						
Bororo	brr	Brazil	16°S	57°W	Borôro	BOR
177.2. Ge-Kaingang						
Apinayé	api	Brazil	5°30S	48°W	Apinayé	APN
Canela	cnl	Brazil	7°S	45°W	Canela	RAM
Canela-Krahô	ckr	Brazil	6°S	45°W	2 varieties: Canela/Krahô	RAM/XRA
Kaingang	kng	Brazil	26°S	52°W	Kaingáng	KGP
Kayapó	kyp	Brazil	9°S	52°W	Kayapó	TXU
Krahô	krh	Brazil	8°S	48°W	Krahô	XRA
Xavánte	xav	Brazil	15°S	52°30W	Xavánte	XAV
Xerénte	xer	Brazil	10°S	48°10W	Xerénte	XER
Xokleng	xok	Brazil	24°30S	50°W	Xokleng	XOK
177.3. Guató						
Guató	gto	Brazil	17°S	58°W	Guató	GTA
177.4. Karajá						
Javaé	jva	Brazil	10°S	50°20W	dialect of Karajá	dKPJ
177.5. Kariri[88]						
Kariri	kri	Brazil	10°S	37°W	Kariri-Xocó	KZW
177.6. Maxakalí						
Maxakalí	max	Brazil	18°S	40°W	Maxakalí	MBL
177.7. Rikbaktsa						
Rikbaktsa	rik	Brazil	11°50S	57°30W	Rikbaktsa	ART
177.8. Yatê						
Fulniô	ful	Brazil	8°S	37°30W	Fulniô	FUN
178. TRUMAI						
Trumai	tru	Brazil	12°S	53°W	Trumaí	TPY
179. NAMBIKUARAN						
Nambikuára	nmb	Brazil	13°S	59°W	Nambikuára (Southern)	NAB
180. KWAZA						
Kwazá	kwz	Brazil	12°35S	60°40W	—	—
181. JABUTÍ						
Jabutí	jbt	Brazil	12°15S	62°15W	Jabutí	JBT
182. CHAPACURA-WANHAN						
Wari’	war	Brazil	11°20S	65°W	Pakaásnovos	PAV
183. MURA						
Pirahã	prh	Brazil	7°S	62°W	Múra-Pirahã	MYP

[88] Listed as Unclassified by *Ethnologue*.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
184. ARAUAN						
Banawá	bna	Brazil	7°S	65°W	Banawá	BNH
Culina	cul	Brazil, Peru	6°S	70°30W	Culina	CUL
Dení	den	Brazil	6°S	67°W	Dení	DAN
Jamamadi	jmm	Brazil	7°30S	67°W	Jamamadi	JAA
Jarawara	jwr	Brazil	7°30S	65°30W	Jaruára	JAP
Paumarí	pau	Brazil	6°S	64°W	Paumarí	PAD
185. KATUKINAN						
Canamarí	cnm	Brazil	6°30S	68°30W	Kanamarí	KNM
186. TACANAN						
Araona	ana	Bolivia	12°20S	67°45W	Araona	ARO
Cavineña	cav	Bolivia	13°20S	66°30W	Cavineña	CAV
Ese Ejja	ese	Bolivia, Peru	11°S	66°W	Ese Ejja	ESE
Tacana	tac	Bolivia	13°30S	68°W	Tacana	TNA
187. ITONAMA						
Itonama	ito	Bolivia	12°50S	64°20W	Itonama	ITO
188. CAYUVAVA						
Cayuvava	cyv	Bolivia	13°30S	65°30W	Cayubaba	CAT
189. MOVIMA						
Movima	mov	Bolivia	13°50S	65°40W	Movima	MZP
190. MOSETENAN						
Mosetén	mos	Bolivia	14°45S	67°50W	Tsimané	CAS
191. YURACARE						
Yuracare	yrc	Bolivia	16°35S	65°15W	Yuracare	YUE
192. PUQUINA						
Puquina	puq	Bolivia, Peru	16°30S	68°30W	—	—
193. URU-CHIPAYA						
Uru	uru	Bolivia	16°45S	69°W	Uru	URE
194. KUNZA						
Atacameño	atm	Chile, Bolivia, Argentina	23°S	69°W	Kunza	KUZ
195. CHIQUITO[89]						
Chiquito	cqt	Bolivia	17°30S	60°W	Chiquitano	CAX
196. ZAMUCOAN						
Ayoreo	ayr	Paraguay	20°15S	59°15W	Ayoreo	AYO
197. MASCOIAN						
Guana	gna	Paraguay	22°S	58°W	Guana	GVA
Lengua	lng	Paraguay	22°30S	59°W	Lengua	LEG
198. MATACOAN[90]						
Chorote	crt	Argentina, Bolivia, Paraguay	22°30S	62°30W	2 varieties: Chorote (Iyojwa'ja)/ Chorote (Iyo'wujwa)	CRT/CRQ

[89] Treated by *Ethnologue* as belonging to Macro-Ge.

[90] Treated by *Ethnologue* as one branch of Mataco-Guaicuru, which is not recognized here as a family.

	Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
	Chulupí	chu	Paraguay	23°30S	60°30W	Chulupí	CAG
	Maca	mca	Paraguay	25°S	57°30W	Maca	MCA
	Wichí	wch	Argentina, Bolivia	22°30S	62°35W	3 varieties: Wichí Lhamtés Güisnay/. . .	MZH/. . .
199. GUAICURUAN[91]							
	Abipón	abi	Argentina	29°S	61°W	—	—
	Kadiwéu	kdw	Brazil	20°S	57°W	Kadiwéu	KBC
	Mocoví	mcv	Argentina	27°S	62°W	Mocoví	MOC
	Pilagá	pga	Argentina	25°S	60°W	Pilagá	PLG
	Toba	tob	Argentina	26°30S	59°W	Toba	TOB
200. LULE-VILELA							
	Lule	lul	Argentina	28°S	64°W	—	—
201. HUARPE							
	Allentiac	alc	Argentina	31°S	68°W	—	—
202. ARAUCANIAN							
	Mapudungun	map	Chile	38°S	72°W	Mapudungun	ARU
203. ALACALUFAN							
	Qawasqar	qaw	Chile	49°S	75°W	Qawasqar	ALC
204. CHON							
204.1. Chon Proper							
	Selknam	sel	Argentina	53°S	70°W	Ona	ONA
	Tehuelche	teh	Argentina	48°S	68°W	Tehuelche	TEH
204.2. Puelche[92]							
	Gününa Küne	gku	Argentina	41°S	67°W	Puelche	PUE
205. YÁMANA							
	Yahgan	yah	Chile	55°S	68°W	Yámana	YAG
206. CREOLES[93] ***AND PIDGINS***							
	Angolar	ago	São Tomé e Príncipe	0°15N	6°30E	Angolar	AOA
	Berbice Dutch Creole	bdc	Guyana	5°20N	58°W	Dutch (Berbice Creole)	BRC
	Bislama	bsm	Vanuatu	16°S	168°E	Bislama	BCY
	Broken	bro	Australia	10°10S	143°E	Torres Strait Creole	TCS
	Chavacano	cvc	Philippines	7°N	122°05E	Chavacano	CBK
	Guadeloupe Creole	gdl	Guadeloupe	16°15N	61°30W	dialect of Lesser Antillean Creole French	dDOM
	Guianese French Creole	gfr	French Guiana	5°N	53°W	French Guianese Creole French	FRE
	Guinea Bissau Crioulo	gbc	Guinea-Bissau	12°N	15°W	Crioulo (Upper Guinea)	POV
	Haitian Creole	hcr	Haiti	19°N	72°30W	French (Haitian Creole)	HAT
	Hawaiian Creole	hwc	United States	21°30N	158°W	Hawai'i Creole English	HAW
	Jamaican Creole	jcr	Jamaica	18°10N	77°15W	dialect of Southwestern Carriban Creole English	dJAM
	Kituba	ktb	Democratic Republic of Congo	5°S	17°30E	Kituba	KTU
	Kriol (Fitzroy Crossing)	kfc	Australia	18°10S	125°35E	—	—
	Kriol (Ngukurr)	knq	Australia	14°50S	135°E	dialect of Kriol	dROP

[91] Treated by *Ethnologue* as one branch of Mataco-Guaicuru, which is not recognized here as a family.
[92] Gününa Küne is treated as a language isolate by *Ethnologue*.
[93] The notion of language family does not apply to creoles, since they have genealogical roots in at least two languages, generally from two separate language families.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Lesser Antillean French Creole	lcr	St. Lucia, Dominica, Grenada, Guadeloupe, Guyana, Martinique, Trinidad and Tobago	15°N	61°10W	Lesser Antillean Creole French	DOM
Malacca Creole	mlc	Malaysia	2°S	129°E	Malaccan Creole Portuguese	MCM
Martinique Creole	mqc	Martinique	14°40N	61°W	dialect of Lesser Antillean Creole French	dDOM
Mauritian Creole	mcr	Mauritius	20°20S	57°30E	Morisyen	MFE
Melayu Betawi	meb	Indonesia	6°30S	107°E	Betawi	BEW
Mískito Coast English Creole	mce	Belize	16°15N	88°50W	dialect of Northern Central America Creole	dBZI
Naga Pidgin	npn	India	25°40N	94°E	Naga Pidgin	NAG
Ndyuka	ndy	Suriname, French Guiana	5°N	54°30W	Aukan	DJK
Nigerian Pidgin	npi	Nigeria	6°30N	3°20E	Pidgin (Nigerian)	PCM
Nubi	nub	Uganda	0°40N	32°05E	Nubi	KCN
Papiamentu	pap	Netherlands Antilles	12°15N	69°W	Papiamentu	PAE
Príncipense	pri	São Tomé e Príncipe	1°37N	7°22E	Principense	PRE
Russian-Chinese Pidgin (Birobidjan)	rcp	Russia	48°48N	132°57E	—	—
Saramaccan	srm	Suriname, French Guiana	4°30N	55°30W	Saramaccan	SRM
Seychelles Creole	sey	Seychelles	4°45S	55°30E	Seselwa Creole French	CRS
Sranan	sra	Suriname	5°50N	55°20W	Sranan	SRN
Tayo	tay	New Caledonia	22°S	166°20E	Tayo	CKS
Tok Pisin	tpi	Papua New Guinea	9°30S	147°10E	Tok Pisin	PDG

207. SIGN LANGUAGES[94]

Adamorobe Sign Language	ada	Ghana	5°50N	0°10W	Adamorobe Sign Language	ADS
American Sign Language	asl	United States, Canada	39°N	78°W	American Sign Language	ASE
Auslan	aus	Australia	30°S	145°E	Australian Sign Language	ASF
British Sign Language	bsl	United Kingdom	52°N	1°W	British Sign Language	BHO
Chinese Sign Language	csl	China	35°N	115°E	Chinese Sign Language	CSL
Deutsche Gebärdensprache	dge	Germany	52°N	11°E	German Sign Language	GSG
Finnish Sign Language	fsl	Finland	62°N	26°E	Finnish Sign Language	FSE
Greek Sign Language	gsl	Greece	38°30N	22°E	Greek Sign Language	GSS
Hong Kong Sign Language	hks	China	22°30N	114°E	—	—
Icelandic Sign Language	ics	Iceland	65°N	18°W	Icelandic Sign Language	ICL
Indo-Pakistani Sign Language (Indian dialects)	ipi	India	24°N	80°E	Indian Sign Language	INS
Indo-Pakistani Sign Language (Karachi dialect)	ipk	Pakistan	28°N	68°E	dialect of Pakistan Sign Language	dPKS
International Sign	isl		40°N	170°W	—	—
Irish Sign Language	irs	Ireland	53°30N	7°30W	Irish Sign Language	ISG
Israeli Sign Language	iss	Israel	32°N	34°50E	Israeli Sign Language	ISL
Kata Kolok	kkl	Indonesia	8°15S	115°10E	Bali Sign Language	BQY
Kenyan Sign Language	ksl	Kenya	0°N	38°E	Kenyan Sign Language	XKI
Langue des Signes Française	lsf	France	47°N	3°E	French Sign Language	FSL
Langue des Signes Québecoise	lsq	Canada	48°N	75°W	Quebec Sign Language	FCS
Lengua de Señas Argentina	lsa	Argentina	34°S	63°W	Argentine Sign Language	AED
Lengua de Señas Española	lse	Spain	40°N	3°W	Spanish Sign Language	SSP
Língua de Sinais Brasileira	lsb	Brazil	15°S	48°W	Brazilian Sign Language	BZS
Língua Gestual Portuguesa	lgp	Portugal	40°N	8°W	Portuguese Sign Language	PSR
Lingua Italiana dei Segni	lii	Italy	42°N	13°E	Italian Sign Language	ISE
Lughat al-Isharat al-Lubnaniya	lgh	Lebanon	34°N	35°50E	—	—
Nederlandse Gebarentaal	ned	Netherlands	52°15N	5°30E	Dutch Sign Language	DSE
New Zealand Sign Language	nzs	New Zealand	43°S	172°E	New Zealand Sign Language	NZS
Nihon Shuwa (Japanese Sign Language)	nih	Japan	36°N	139°E	Japanese Sign Language	JSL
Norsk Tegnspråk	nte	Norway	61°N	9°E	Norwegian Sign Language	NSL
Plains-Indians Sign	psl	United States	40°N	100°W	Plains Indian Sign Language	PSD

[94] It is clear that since deaf sign languages have arisen independently in different parts of the world, there must be a number of different language families of sign languages. However, we do not attempt to offer a genealogical classification here.

Language	*WALS* code	Country	Lat	Long	*Ethnologue* Name	*Ethn.* code
Language						
Russian Sign Language	rsl	Russia	56°N	44°E	Russian Sign Language	RSL
South Korean Sign Language	ssl	South Korea	37°N	128°E	Korean Sign Language	KVK
Svenska Teckenspråket	svt	Sweden	59°N	15°E	Swedish Sign Language	SWL
Taiwanese Sign Language (Ziran Shouyu)	tzi	Taiwan	24°N	121°E	Taiwanese Sign Language	TSS
Tanzania Sign Language	tsl	Tanzania	6°S	35°E	Tanzanian Sign Language	TZA
Thai Sign Language	ths	Thailand	15°N	101°E	Thai Sign Language	TSQ
Türk Isaret Dili	tui	Turkey	39°N	34°E	Turkish Sign Language	TSM
Ugandan Sign Language	ugs	Uganda	1°N	32°30E	Ugandian Sign Language	UGN
Urubú Sign Language	usl	Brazil	2°20S	46°W	Urubú-Kaapor Sign Language	UKS
Vlaamse Gebarentaal	vla	Belgium	51°N	4°E	dialect of Belgian Sign Language	dBVS

References

A

Abbott, Miriam. 1991. 'Macushi'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages* 3, pp. 23–160. Berlin: Mouton de Gruyter.

Abdel-Hafiz, Ahmed Sokarno. 1988. 'A Reference Grammar of Kunuz Nubian'. State University of New York at Buffalo Dissertation.

Abraham, P. T. 1985. *Apatani Grammar*. Manasagangotri, Mysore: Central Institute of Indian Languages.

Adelaar, Willem F. H., with Pieter C. Muysken. 2004. *The Languages of the Andes*. Cambridge: Cambridge University Press.

Adwiraah, Eleonore. 1989. *Grammatik des Gworok (Kagoro): Phonologie, Tonologie, Morphologie und Textanalyse*. Europäische Hochschulschriften xxi, linguistik 71. Frankfurt: Peter Lany.

Agha, Asif. 1994. 'Honorification'. *Annual Review of Anthropology* 23: 277–302.

Agrell, Sigurd. 1908. *Aspektänderung und Aktionsartbildung beim polnischen Zeitworte: ein Beitrag zum Studium der indogermanischen Präverbia und ihrer Bedeutungsfunktionen*. Lunds Universitets Årsskrift. Lund: Lunds Universitet.

Aikhenvald, Alexandra Y. 1990. *Sovremennyj ivrit*. Moscow: Nauka.

—— 1994. 'Classifiers in Tariana'. *Anthropological Linguistics* 36: 407–65.

—— 1996. *Bare*. Languages of the World / Materials 100. München: Lincom Europa.

—— 1998. 'Warekena'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages* 4, pp. 225–439. Berlin and New York: Mouton de Gruyter.

—— 1999. 'Arawak'. In R. M. W. Dixon and Alexandra Y. Aikhenvald, eds., *The Amazonian Languages*, pp. 3–104. Cambridge: Cambridge University Press.

—— 2000. *Classifiers: a Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

—— 2005. *Evidentiality*. Oxford: Oxford University Press.

—— (forthcoming a). 'Reciprocal and Sociative in Tariana: their Genetic and Areal Properties'. In Nedjalkov et al., *Typology of Reciprocal Constructions*. München: Lincom Europa.

—— (forthcoming b). 'Reciprocals and Reflexives in North-Arawak Languages of the Upper Rio Negro (Warekena, Bare, Baniwa of Içana)'. In Nedjalkov et al., *Typology of Reciprocal Constructions*. München: Lincom Europa.

—— and R. M. W. Dixon. 1998. 'Evidentials and Areal Typology: A Case Study from Amazonia'. *Language Sciences* 20.3: 241–57.

—— —— (eds.) 2003. *Studies in Evidentiality*. Typological Studies in Language 54. Amsterdam: Benjamins.

Aissen, Judith. 2003. Differential Object Marking: Iconicity versus Economy. *Natural Language and Linguistic Theory* 21.3: 435–83.

Aksu-Koç, Ayhan A., and Dan I. Slobin. 1986. 'A Psychological Account of the Development and Use of Evidentials in Turkish'. In Wallace Chafe, and Johanna Nichols, eds., *Evidentiality: the Linguistic Coding of Epistemology*, pp. 159–67. Norwood, NJ: Ablex.

Alekseev, Mikhail E. 1994. 'Rutul'. In Rieks Smeets, ed., *The Indigenous Languages of the Caucasus* 4, Part 2, pp. 213–58. Delmar, New York: Caravan Books.

Alexander, Ruth Mary. 1988. 'A Syntactic Sketch of Ocotepec Mixtec'. In C. Henry Bradley and Barbara E. Hollenbach, eds., *Studies in the Syntax of Mixtecan Languages* 1, pp. 151–304. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Aliroev, Ibragim Junusovich. 1997. *Chechenskij jazyk*. Moscow: Academia.

Allan, Robin, Philip Holmes, and Tom Lundskær-Nielsen. 1995. *Danish: A Comprehensive Grammar*. London: Routledge.

Alpatov, Vladimir M., Anna Ju. Bugaeva, and Vladimir P. Nedjalkov. (forthcoming). 'Reciprocals and Sociatives in Ainu'. In Nedjalkov et al., *Typology of Reciprocal Constructions*. München: Lincom Europa.

—— and Vladimir P. Nedjalkov. (forthcoming). 'Reciprocal, Sociative and Competitive Constructions in Japanese'. In Nedjalkov et al., *Typology of Reciprocal Constructions*. München: Lincom Europa.

—— and Vera Podlesskaya. 1995. 'Converbs in Japanese'. In M. Haspelmath and E. König, eds., *Converbs in Cross-Linguistic Perspective*, pp. 465–85. Berlin: Mouton de Gruyter.

Alsina, Alex. 1996. 'Passive Types and the Theory of Object Asymmetries'. *Natural Language and Linguistic Theory* 14: 673–723.

—— and Sam Mchombo. 1993. 'Object Asymmetries and the Chichewa Applicative Construction'. In Sam Mchombo, ed., *Theoretical Aspects of Bantu Grammar*, pp. 17–45. Stanford: Center for the Study of Language and Information.

Ambrazas, Vytautas, Emma Geniušienė, Aleksas Girdenis, Nijolė Sližienė, Dalija Tekorienė, Adelė Valeckienė, and Elena Valiulytė. 1997. *Lithuanian Grammar*. Vilnius: Baltos Lankos.

Ammann, Andreas and Johan van der Auwera. 2004. 'Complementizer-headed Main Clauses for Volitional Moods in the Languages of South-Eastern Europe: A Balkanism?'. In Olga Mišeska Tomić, ed., *Balkan Syntax and Semantics* pp. 293–314. Amsterdam: John Benjamins.

Andersen, Paul Kent. 1983. *Word Order Typology and Comparative Constructions*. Amsterdam: John Benjamins.

Andersen, Torben. 1988. 'Ergativity in Parï, a Nilotic OVS language'. *Lingua* 75: 289–324.

Anderson, Gregory D. S. 2003a. 'Towards a Phonological Typology of Native Siberia'. In Dee Ann Holisky and Kevin J. Tuite, eds., *Current Trends in Caucasian, East European and Inner Asian Linguistics, Papers in Honor of Howard I. Aronson*, pp. 1–22. Amsterdam: John Benjamins.

—— 2003b. 'Yeniseic Languages from a Siberian Areal Perspective'. In E. J. Vajda and G. D. S. Anderson, eds., *Studia Yeniseica: in Honor of Heinrich Werner*, Special issue of *Sprachtypologie und Universalienforschung* 56.1/2: 12–39. Berlin: Akademie Verlag.

Anderson, John, and Edward Keenan. 1985. 'Deixis'. In Timothy Shopen, ed., *Language Typology and Syntactic Description* 3, pp. 259–308. Cambridge: Cambridge University Press.

Anderson, Stephen. 1985. 'Inflectional Morphology'. In Timothy Shopen, ed., *Language Typology and Syntactic Description* 3, pp. 151–201. Cambridge: Cambridge University Press.

Andrade, Manuel J. 1933. 'Quileute'. In Franz Boas, ed., *Handbook of American Indian Languages* 3, Bureau of American Ethnology, Bulletin 40.3, pp. 151–292. New York: Colombia University Press.

Andronov, Mikhail S. 1980. *The Brahui Language*. Moscow: Nauka.

Anonymous. 1985. *Miskitu Bila Aisanka: Gramatica Miskita*. Managua: Centro de Investigaciones y Documentación de la Costa Atlántica.

Aoki, Haruo. 1963. 'Reduplication in Nez Perce'. *International Journal of American Linguistics* 29: 42–4.

—— 1970. *Nez Perce Grammar*. UCPL 62. Berkeley and Los Angeles: University of California Press.

—— 1994. *Nez Perce Dictionary*. UCPL 122. Berkeley and Los Angeles: University of California Press.

Ariel, Mira. 2000. 'The Development of Person Agreement Markers: from Pronouns to Higher Accessibility Markers'. In M. Barlow and S. Kemmer, eds., *Usage-based Models of Language*, pp. 197–260. Stanford, CA: CSLI Publications.

Armbruster, Charles Herbert. 1960. *Dongolese Nubian: A Grammar*. Cambridge: Cambridge University Press.

Arnott, D. W. 1967. 'Some Reflections on the Content of Individual Classes in Fula and Tiv'. In André Martinet et al., eds., *La classification nominale dans les langues négro-africaines*, Colloques internationaux du Centre National de la Recherche Scientifique—Sciences humaines, Aix-en-Provence, 3–7 juillet 1967, pp. 45–74. Paris: Éditions du Centre national de la recherche scientifique.

Arnott, D. W. 1970. *The Nominal and Verbal Systems of Fula*. Oxford: Clarendon Press.

Aronoff, Mark 1994. *Morphology by Itself: Stems and Inflectional Classes*. Linguistic Inquiry Monograph 22. Cambridge, MA: MIT Press.

Asher, R. E. 1982. *Tamil*. Amsterdam: North-Holland.

—— 1985. *Tamil*. London: Croom Helm.

—— and T. C. Kumari. 1997. *Malayalam*. London: Routledge.

Ashton, E. O. 1947. *Swahili Grammar*. London: Longmans, Green and Camp & Co. [2nd edn.]

—— et al. 1954. *Luganda Grammar*. London: Longmans, Green & Co.

Aski, Janice. 1995. 'Verbal Suppletion: an Analysis of Italian, French, and Spanish to go'. *Linguistics* 33: 403–32.

Austin, Peter. 1981. *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

Austin, Peter. 1997. 'Causatives and Applicatives in Australian Aboriginal Languages'. In Kazuo Matsumura and Tooru Hayasi, eds., *The Dative and Related Phenomena*, pp. 165–225. Tokyo: Hitsuji Shobo.
Awobuluyi, A. Oladele. 1978. *Essentials of Yoruba Grammar*. Ibadan: Oxford University Press Nigeria.

B

Babby, Leonard H. 1987. 'Case, Prequantifiers, and Discontinuous Agreement in Russian'. *Natural Language and Linguistic Theory* 5: 91–138.
Bach, Emmon. 1996. 'A Note on Quantification and Blankets in Haisla'. In E. Bach, E. Jelinek, A. Kratzer, and B. H. Partee, eds., *Quantification in Natural Languages*, pp. 13–20. Dordrecht: Kluwer Academic Publishers.
Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
—— 1998. 'On the Structural Positions of Themes and Goals'. In Johan Rooryck and Laurie Zaring, eds., *Phrase Structure and the Lexicon*, pp. 7–34. Dordrecht: Kluwer.
Bánhidi, Zoltán, Zoltán Jókay, and Dénes Szabó. 1975. *Lehrbuch der ungarischen Sprache*. München: Hueber.
Barnes, Janet. 1984. Evidentials in the Tuyuca Verb. *International Journal of American Linguistics* 50: 255–71.
Bartens, Hans-Hermann. 1989. *Lehrbuch der saamischen (lappischen) Sprache*. Hamburg: Helmut Buske.
Barulin, Aleksandr. 1980. 'Kategorija čisla v mestoimenijax'. [The category of number in pronouns]. In V. M. Andrjušenko, ed., *Issledovanija v oblasti grammatiki i tipologii jazykov*. Moscow: MGU.
Bauer, Winifred. 1993. *Maori*. London and New York: Routledge.
Beaumont, Clive H. 1979. *The Tigak Language of New Ireland*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
Beckwith, Christopher I. 1992. 'Classifiers in Hungarian'. In I. Kenesei and C. Pléh, eds., *Approaches to Hungarian, Volume 4: The Structure of Hungarian*, pp. 197–206. Szeged: JATE.
Beddor, P. 1993. 'The Perception of Nasal Vowels'. In M. Huffman and R. Krakow, eds., *Nasals, Nasalization and the Velum*, pp. 171–96. San Diego: Academic Press.
Beller, Richard and Patricia Beller. 1977. 'Huasteca Nahuatl'. In Ronald W. Langacker, ed., *Modern Aztec Grammatical Sketches*, Studies in Uto-Aztecan Grammar 2, pp. 199–306. Dallas: Summer Institute of Linguistics.
Bender, M. Lionel. 1996. *Kunama*. München: Lincom Europa.
—— 1997. *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom Europa. [2nd edn.]
Bendor-Samuel, David. 1972. *Hierarchical Structures in Guajajara*. Norman, Oklahoma: Summer Institute of Linguistics.
Benjamin, Geoffrey. 1976. *An Outline of Temiar Grammar*. Austroasiatic Studies 1. Oceanic Linguistics Special Publication. Honolulu: University Press of Hawaii.
Benveniste, Émile. 1966a. 'Structure des relations de personne dans le verbe'. In Emile Benveniste, *Problèmes de linguistique générale*, pp. 225–36. Paris: Gallimard.
—— 1966b. 'La phrase nominale.' In Émile Benveniste, *Problèmes de linguistique générale*, pp. 151–67. Paris: Gallimard.
—— 1971. *Problems in General Linguistics*. Translated by Mary Elizabeth Meek. Coral Gables, Fa: University of Miami Press.
Berchem, Jörg. 1991. *Referenzgrammatik des Somali*. Köln: Omimee.
Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager*. Wiesbaden: Harrassowitz.
Berlin, Brent and Paul Kay. 1969. *Basic Color Terms: their Universality and Evolution*. Berkeley and Los Angeles: University of California Press.
Bernini, Giuliano and Paolo Ramat. 1996. *Negative Sentences in the Languages of Europe: A Typological Approach*. Berlin: Mouton de Gruyter.
Berry, Keith and Christine Berry. 1999. *A Description of Abun: A West Papuan Language of Irian Jaya*. Pacific Linguistics, Series B-115. Canberra: Australian National University.
Besha, Ruth M. 1993. *A Classified Vocabulary of the Shambala Language, with Outline Grammar*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
Besnier, Niko. 2000. *Tuvaluan, a Polynesian Language of the Central Pacific*. London and New York: Routledge.
Bhaskararao, Peri. 1998. 'Gadaba'. In Sanford B. Steever, ed., *The Dravidian Languages*, pp. 328–55. London: Routledge.
Bhat, D. N. S. 2000. 'The Indefinite-Interrogative Puzzle'. *Linguistic Typology* 4.3: 365–400.
—— 2004. *Pronouns*. Oxford: Oxford University Press.
—— and M. S. Ningomba. 1997. *Manipuri Grammar*. München: Lincom Europa.
Bhatia, Tej K. 1993. *Punjabi: A Cognitive-Descriptive Grammar*. Routledge Descriptive Grammar Series. London: Routledge.
Bickel, Balthasar. 2003. 'Prosodic Tautomorphemicity in Sino-Tibetan'. In David Bradley, Randy J. LaPolla, Boyd Michailovsky, and Graham Thurgood, eds., *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honor of James A. Matisoff*. Canberra: Pacific Linguistics.
—— and Johanna Nichols. 2002. 'Autotypologizing Databases and their Use in Fieldwork'. In Peter Austin, Helen Dry, and Peter Wittenburg, eds., *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas, 25–26 May 2002*. Nijmegen: Max Planck Institute for Psycholinguistics.
—— —— 2003. *Typological Enclaves*. Paper presented at the 5th Biennial Meeting of the Association for Linguistics Typology, Cagliari, September 2003.
—— —— 2005. 'Inflectional Morphology'. In Shopen, Timothy, ed., *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press. [2nd edition].
Bielec, Dana. 1998. *Polish: An Essential Grammar*. London: Routledge.
Biggs, Bruce. 1969. *Let's Learn Maori*. Wellington: Reed.
Biligiri, Hemmige S. 1965. *Kharia: Phonology, Grammar and Vocabulary*. Pune: Deccan College.
Bird, C. and M. Kante. 1976. *An Kan Bamanakan Kalan: Intermediate Bambara*. Bloomington: Indiana University Linguistics Club.
Birk, D. B. W. 1976. *The Malakmalak Language, Daly River (Western Arnhem Land)*. Pacific Linguistics, Series B-45. Canberra: Australian National University.
Bittner, Maria. 1987. 'On the Semantics of the Greenlandic Antipassive and Related Constructions'. *International Journal of American Linguistics* 53: 194–231.
—— and Kenneth Hale. 1996. 'The Structural Determination of Case and Agreement'. *Linguistic Inquiry* 27: 1–68.
Blackings, Mairi, and Nigel Fabb. 2003. *A Grammar of Ma'di*. Berlin: Mouton de Gruyter.
Black-Rogers, Mary B. 1982. 'Algonquian Gender Revisited: Animate Nouns and Ojibwa "Power"—an Impasse?' *Papers in Linguistics* 15: 59–76.
Blake, Barry. 1994. *Case*. Cambridge: Cambridge University Press.
Blevins, Juliette. 2001. *Nhanda: an Aboriginal Language of Western Australia*. Honolulu: University of Hawaii Press.
Bloomfield, Leonard. 1957. *Eastern Ojibwa: Grammatical Sketch, Texts and Word List*. Ann Arbor: University of Michigan Press.
—— 1962. *The Menomini Language*. New Haven: Yale University Press.
Boas, Franz. 1911. 'Tsimshian'. In Franz Boas, ed., *Handbook of American Indian Languages 1*, Smithsonian Institution Bureau of American Ethnology Bulletin 40, pp. 283–422. Washington, D.C.: Government Printing Office.
—— and Ella Deloria. 1941. *Dakota Grammar*. Memoirs of the National Academy of Sciences 23, 2nd Memoir. Washington: U.S. Government Printing Office.
Bodding, P. O. 1952. *A Santali Grammar for Beginners*. Benagaria: Santal Mission of the Northern Churches Santali.
Bodomo, Adams. 1997. *The Structure of Dagaare*. Stanford: CSLI Publications.
Boeder, Winfried. 2000. 'Evidentiality in Georgian'. In Lars Johanson and Bo Utas, eds., *Evidentials: Turkic, Iranian and Neighbouring Languages*, Empirical Approaches to Language Typology 24 pp. 275–328. Berlin: Mouton de Gruyter.
Boelaars, J. H. M. C. 1950. *The Linguistic Position of South-Western New Guinea*. Leiden: E. J. Brill.
Bohnhoff, Lee E. 1986. 'Yag Dii (Duru) Pronouns'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 103–30. Tübingen: Gunter Narr.
Bolton, Rosemary Ann. 1990. 'A Preliminary Description of Nuaulu Phonology and Grammar'. University of Texas at Arlington M.A. Thesis.
Bonvillain, Nancy. 1973. *A Grammar of Akwesasne Mohawk*. Ethnology Division, Paper 8. Ottawa: National Museum of Man.
Borgman, Donald M. 1990. 'Sanuma'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages* 2, pp. 15–248. Berlin: Mouton de Gruyter.
Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.
Bossong, Georg. 1998. 'Le marquage différentiel de l'objet dans les langues de l'Europe'. In Jack Feuillet, ed., *Actance et valence dans les langues de l'Europe*, Empirical Approaches to Language Typology 20–2, pp. 193–258. Berlin: Mouton de Gruyter.

Botne, Robert. 1995. 'The Pronominal Origin of an Evidential'. *Diachronica* 12: 201–21.

Bowden, John. 1997. 'Taba (Makian Dalam): Description of an Austronesian language from Eastern Indonesia'. University of Melbourne Ph.D. Dissertation.

—— 2001. *Taba: Description of a South Halmahera Language*. Pacific Linguistics 521. Canberra: Australian National University.

Bowe, Heather. 1990. *Categories, Constituents, and Constituent Order in Pitjantjatjara, an Aboriginal Language of Australia*. London: Routledge.

Boyeldieu, Pascal. 1982. *Deux études laal: Moyen-Chari, Tchad*. Berlin: Dietrich Reimer.

—— 1995. 'Le yakoma'. In Raymond Boyd, ed., *Le système verbal dans les langues oubangiennes*, pp. 113–39. München: Lincom Europa.

Boyes Braem, Penny. 1990. *Einführung in die Gebärdensprache und ihre Erforschung*. Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 11. Hamburg: Signum.

Bradley, David. 1997. 'Tibeto-Burman languages and classification'. In David Bradley, ed., *Tibeto-Burman Languages of the Himalayas*, pp. 1–72. Papers in Southeast Asian Linguistics, No. 14. Pacific Linguistics, Series A, No. 86. Canberra: Australian National University.

Braine, Jean Critchfield. 1970. 'Nicobarese Grammar (Car Dialect)'. University of California, Berkeley Ph.D. Dissertation.

Branson, Jan, Don Miller, I Gede Marsaja and I Wayan Negara. 1996. 'Everyone Here Speaks Sign Language Too: A Deaf Village in Bali, Indonesia'. In Ceil Lucas, ed., *Multicultural Aspects of Sociolinguistics in Deaf Communities*, pp. 39–57. Washington, D.C.: Gallaudet University Press.

Bray, Danys De. 1986. *The Brahui Language 1*, Calcutta: Superintendent Govt. Printing.

Breedveld, J. O. 1995. *Form and Meaning in Fulfulde: A Morphophonological Study of Maasinankoore*. Leiden: Research School CNWS.

Breen, J. Gavin. 1981. 'Margany and Gunya'. In R. M. W. Dixon and B. J. Blake, eds., *Handbook of Australian Languages 2*, pp. 275–393. Canberra: ANU Press.

Breeze, Mary J. 1986. 'Personal Pronouns in Gimira'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 47–70. Tübingen: Gunter Narr.

—— 1990. 'A Sketch of the Phonology and Grammar of Gimira (Benchnon)'. In Richard J. Hayward, ed., *Omotic Language Studies*, pp. 1–67. London: School of Oriental and African Studies, University of London.

Bright, William. 1957. *The Karok Language*. University of California Publications in Linguistics 13. Berkeley: University of California Press.

—— 1965a. 'Luiseño phonemics'. *International Journal of American Linguistics* 31.4: 342–5.

—— 1965b. 'The History of the Cahuilla Sound System'. *International Journal of American Linguistics* 31.3: 241–4.

Bril, Isabelle. 1999. 'Negation in Nêlêmwa (Austronesian, Oceanic Language of Northern New-Caledonia)'. In Even Hovdhaugen and Ulrike Mosel, eds., *Negation in Oceanic Languages, Typological Studies*, pp. 80–95. München: Lincom Europa.

—— (forthcoming). 'Polysemy of the Reciprocal Marker in Nêlêmwa'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

Broadbent, Sylvia M. 1964. *The Southern Sierra Miwok Language*. University of California Publications in Linguistics 38. Berkeley: University of California Press.

Bromley, H. Myron. 1961. *The Phonology of Lower Grand Valley Dani: A Comparative Structural Study of Skewed Phonemic Patterns*. 's-Gravenhage: Martinus Nijhoff.

—— 1981. *A Grammar of Lower Grand Valley Dani*. Pacific Linguistics, Series C-63. Canberra: Australian National University.

Brown, Cecil H. and Witkowski, Stanley R. 1981. 'Figurative Language in a Universalist Perspective'. *American Ethnologist* 8: 596–615.

Brown, Lea. 2001. 'A Grammar of Nias Selatan'. University of Sydney Ph.D. Dissertation.

—— (forthcoming). 'Nias'. In Nikolaus P. Himmelmann and K. Alexander Adelaar, eds., *The Austronesian Languages of Asia and Madagascar*. London and Richmond: Curzon Press.

Brown, Penelope and Stephen C. Levinson. 1987. *Politeness*. Cambridge: Cambridge University Press.

Brown, R. and A. Gilman. 1960. 'The Pronouns of Power and Solidarity'. In T. A. Sebeok, ed., *Style in Language*, pp. 253–76. Cambridge, MA: MIT Press.

Bruce, Les. 1984. *The Alamblak Language of Papua New Guinea (East Sepik)*. Pacific Linguistics, Series C-81. Canberra: Australian National University.

Buddruss, Georg. 1967. *Die Sprache von Sau in Ostafghanistan*. München: J. Kitzinger.

Burling, Robbins. 1965. 'How to Choose a Burmese Numerical Classifier'. In M. E. Spiro, ed., *Context and Meaning in Cultural Anthropology*, pp. 243–64. New York: The Free Press.

Burquest, Donald A. 1986. 'The Pronoun System of some Chadic Languages'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 71–102. Tübingen: Gunter Narr.

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam and Philadelphia: John Benjamins.

—— William Pagliuca, and Revere D. Perkins. 1990. 'On the Asymmetries in the Affixation of Grammatical Material'. In William Croft, Keith Denning and Suzanne Kemmer, eds., *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday*, pp. 1–42. Amsterdam: John Benjamins.

—— —— —— 1991. 'Back to the Future'. In E. C. Traugott and B. Heine, eds., *Approaches to Grammaticalization 2*, pp. 17–58. Amsterdam: John Benjamins.

—— —— —— 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Bystrov, Igor S., and Nonna V. Stankevich. (forthcoming). 'Reciprocal constructions in Vietnamese'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

C

Camargo, Eliane. (forthcoming). 'Reciprocals in Cashinahua'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

Campbell, Carl, and Jody Campbell. 1987. *Yadë Grammar Essentials*. Unpublished manuscript. Summer Institute of Linguistics.

Campbell, Lyle. 1985. *The Pipil Language of El Salvador*. Berlin: Mouton de Gruyter.

Capell, A. and H. E. Hinch. 1970. *Maung Grammar*. Janua Linguarum, Series Practica 98. The Hague: Mouton de Gruyter.

Carlin, Eithne. 1993. *The So Language*. Köln: Institut für Afrikanistik.

Carlson, Robert. 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

—— and Doris Payne. 1989. 'Genitive Classifiers'. In Robert Carlsson, Scott DeLancey, Spike Gildea, Doris Payne, and Anju Saxena, eds., *Proceedings of the Fourth Annual Pacific Linguistics Conference*, pp. 89–119. Eugene: University of Oregon.

Carstairs-McCarthy, Andrew. 1992. *Current Morphology*. London: Routledge.

—— 1994. 'Suppletion'. In S. Asher and F. Simpson, eds., *The Encyclopedia of Language and Linguistics*, pp. 4410–11. Oxford: Pergamon Press.

Caughley, Ross C. 1982. *The Syntax and Morphology of the Verb in Chepang*. Pacific Linguistics, Series B-84. Canberra: Australian National University.

Celo, Pietro. 1996. 'Pragmatic Aspects of the Interrogative Form in Italian Sign Language'. In Ceil Lucas, ed., *Multicultural Aspects of Sociolinguistics in Deaf Communities*, pp. 132–51. Washington, D.C.: Gallaudet University Press.

Chafe, Wallace L. 1996. 'Sketch of Seneca'. In Ives Goddard, ed., *Handbook of American Indians, Vol. 17: Languages*, pp. 551–9. Washington: Smithsonian Institute.

—— and Johanna Nichols. 1986. *Evidentiality: the Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.

Chapman, Shirley, and Desmond C. Derbyshire. 1991. 'Paumari'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 3*, pp. 161–352. Berlin: Mouton de Gruyter.

Charachidze, Georges. 1989. 'Ubykh'. In B. George Hewitt, ed., *The Indigenous Languages of the Caucasus, Volume 2: The Northwest Caucasian Languages*, pp. 357–459. Delmar, N.Y.: Caravan Books.

Chelliah, Shobhana Lakshmi. 1997. *A Grammar of Meithei*. Berlin: Mouton de Gruyter.

Childs, G. Tucker. 1995. *A Grammar of Kisi*. Berlin: Mouton de Gruyter.

Choe, Jae-Woong. 1987. 'Anti-Quantifiers and a Theory of Distributivity'. University of Massachusetts Ph.D. Dissertation.

Chomsky, Noam. 1970. 'Remarks on Nominalization'. In Roderick A. Jacobs and Peter S. Rosenbaum, eds., *Readings in English Transformational Grammar*, pp. 184–221. Boston: Ginn.

—— 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chung, Chul-Hwa and Kyung-Ja Chung. 1996. 'Kuot Grammar Essentials'. In John M. Clifton, ed., *Two Non-Austronesian Grammars from the Islands*, Data Papers on Papua New Guinea Languages 42, pp. 1–75. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Chung, Sandra. 1998. *The Design of Agreement*. Chicago: University of Chicago Press.

Chung, Sandra and Alan Timberlake. 1985. 'Tense, Aspect and Mood'. In Timothy Shopen, ed., *Language Typology and Syntactic Description 3*, pp. 241–58. Cambridge: Cambridge University Press.
Churchward, C. Maxwell. 1953. *Tongan Grammar*. London: Oxford University Press.
Cincius, V. I., ed. 1975. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov*. Leningrad: Nauka.
Clairis, Christos. 1985. *El qawesqar: lingüística fueguina, teoría y descripción*. Estudios Filológicos, Anejo 12. Valdivia: Universidad Austral de Chile.
Clark, Eve V. 1978. 'Locationals: Existential, Locative and Possessive Constructions'. In J. Greenberg, C. Ferguson, and E. Moravcsik, eds., *Universals of Human Language, Volume 4: Syntax*, pp. 86–126. Stanford: Stanford University Press.
Coates, Jennifer. 1983. *The Semantics of Modal Auxiliaries*. London: Croom Helm.
Cohn, A. 1993. 'The Status of Nasalized Continuants'. In M. Huffman and R. Krakow, eds., *Nasals, Nasalization and the Velum*, pp. 329–67. San Diego: Academic Press.
Colarusso, John. 1992. *A Grammar of the Kabardian Language*. Calgary: University of Calgary Press.
Cole, Peter. 1982. *Imbabura Quechua*. Lingua Descriptive Studies 5. Amsterdam: North-Holland.
Collins, B. 1962. *Tonga Grammar*. London: Longmans, Green & Co.
Comrie, Bernard. 1976a. 'The Syntax of Action Nominals: A Cross-Linguistic Study'. *Lingua* 40: 177–201.
—— 1976b. 'The Syntax of Causative Constructions: Cross-Language Similarities and Divergences'. In Masayoshi Shibatani, ed., *Syntax and Semantics, Volume 6: The Grammar of Causative Constructions*, pp. 261–312. New York: Academic Press.
—— 1978. 'Ergativity'. In Winfred P. Lehmann, ed., *Syntactic Typology: Studies in the Phenomenology of Language*, pp. 329–94. Austin: University of Texas Press.
—— 1985. *Tense*. Cambridge: Cambridge University Press.
—— 1989. *Language Universals and Linguistic Typology*. Oxford: Blackwell. [2nd edn.]
—— 1991. 'Form and Function in Identifying Cases'. In Frans Plank, ed., *Paradigms: The Economy of Inflection*, pp. 41–56. Berlin: Mouton de Gruyter.
—— 1997. 'Some Problems in the Theory and Typology of Numeral Systems'. In B. Palek, ed., *Proceedings of LP' 96*, pp. 41–56. Prague: Charles University Press.
—— 1998. 'Rethinking the typology of relative clauses', *Language Design* 1: 59–86.
—— 1999. 'Spatial Cases in Daghestanian Languages'. *Sprachtypologie und Universalienforschung* 52.2: 108–17.
—— and Maria Polinsky. 1998. 'Gender in a Historical Perspective: Radial Categories Meet Language Change'. In Carol F. Justus and Edgar C. Polomé, eds., *Language Change and Typological Variation: In Honor of Winfred P. Lehmann on the Occasion of his 83rd Birthday, Volume 2: Grammatical Universals and Typology*, Journal of Indo-European Studies Monograph 31, pp. 566–89. Washington D.C.: Institute for the Study of Man.
—— and Sandra A. Thompson. 1985. 'Lexical Nominalization'. In Timothy Shopen, ed., *Language Typology and Syntactic Description 3*, pp. 349–98. Cambridge: Cambridge University Press.
Contini-Morava, Ellen. 1989. *Discourse Pragmatics and Semantic Categorization: The Case of Negation and Tense-Aspect with Special Reference to Swahili*. Discourse Perspectives on Grammar 1. Berlin: Mouton de Gruyter.
Conzemius, Eduard. 1929. 'Notes on the Miskito and Sumu Languages of Eastern Nicaragua and Honduras'. *International Journal of American Linguistics* 5: 57–115.
Cooke, Joseph R. 1968. *Pronominal Reference in Thai, Burmese, and Vietnamese*. Berkeley and Los Angeles: University of California Press.
Cooreman, Ann. 1987. *Transitivity and Discourse Continuity in Chamorro Narratives*. Berlin: Mouton de Gruyter.
—— 1988. 'The Antipassive in Chamorro: Variations on the Theme of Transitivity'. In Masayoshi Shibatani, ed., *Passive and Voice*, pp. 561–93. Amsterdam: John Benjamins.
—— 1994. 'A functional typology of antipassives'. In Barbara Fox and Paul Hopper, eds., *Voice: Form and Function*, pp. 49–88. Amsterdam: John Benjamins.
Corbett, Greville G. 1982. 'Gender in Russian: an Account of Gender Specification and its Relationship to Declension'. *Russian Linguistics* 6: 197–232.
—— 1991. *Gender*. Cambridge: Cambridge University Press.
—— 2000. *Number*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
—— and Norman M. Fraser. 2000. 'Gender Assignment: a Typology and a Model'. In Gunter Senft, ed., *Systems of Nominal Classification, Language, Culture and Cognition 4*, pp. 293–325. Cambridge: Cambridge University Press.
—— and Richard J. Hayward. 1987. 'Gender and Number in Bayso'. *Lingua* 73: 1–28.
Cornyn, William. 1944. *Outline of Burmese Grammar*. Language Dissertation 38. Supplement to Language 20, Number 4. Baltimore: Linguistic Society of America.
Couro, Ted, and Margaret Langdon. 1975. *Let's Talk 'Iipay Aa: An Introduction to the Mesa Grande Diegueño Language*. Ramona, California: Ballena Press.
Cowell, Mark W. 1964. *A Reference Grammar of Syrian Arabic*. Washington: Georgetown University Press.
Craig, Colette Grinevald. 1977. *The Structure of Jacaltec*. Austin: University of Texas Press.
Craig, Colette Grinevald, ed. 1986. *Noun Classes and Categorization*. Typological Studies in Language 7. Amsterdam: John Benjamins.
Crazzolara, J. P. 1933. *Outlines of a Nuer Grammar*. Vienna: Anthropos.
—— 1960. *A Study of the Logbara (Ma'di) Language*. London: Oxford University Press.
Creider, Chet A., and Jane Tapsubei Creider. 1989. *A Grammar of Nandi*. Hamburg: Helmut Buske.
Creissels, Denis. 2000. Typology. In Bernd Heine and Derek Nurse, eds., *African Languages: an Introduction*, pp. 231–58. Cambridge: Cambridge University Press.
Cristofaro, Sonia. 1998. Deranking and Balancing in Different Subordination Relations: a Typological Study. *Sprachtypologie und Universalienforschung* 51: 3–42.
—— 2003. *Subordination*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.
Croft, William. 1990. *Typology and Universals*. Cambridge: Cambridge University Press.
Crowell, Thomas Harris. 1979. 'A Grammar of Bororo'. Cornell University Ph.D. Dissertation.
Crowley, Terry. 1982. *The Paamese Language of Vanuatu*. Pacific Linguistics, Series B-87. Canberra: Australian National University.
—— 1983. 'Uradhi'. In R. M. W. Dixon and Barry J. Blake, eds., *Handbook of Australian Languages 3*, pp. 306–428. Amsterdam: John Benjamins.
—— 2002. 'Gela'. In John Lynch, Malcolm Ross, and Terry Crowley, eds., *The Oceanic Languages*, pp. 525–37. Richmond: Curzon.
Curnow, Timothy J. 1997. 'A Grammar of Awa Pit (Cuaiquer): An Indigenous Language of South-Western Colombia'. Australian National University Ph.D. Thesis.
Cutler, Anne, John A. Hawkins, and Gary Gilligan. 1986. 'The Suffixing Preference: a Processing Explanation'. *Linguistics* 23: 723–58.
Cyffer, Norbert. 1974. *Syntax des Kanuri*. Hamburg: Helmut Buske.
—— and John P. Hutchison, eds. 1990. *Dictionary of the Kanuri Language*. Dordrecht: Foris.
Cysouw, Michael, 2002. 'Interpreting Typological Clusters'. *Linguistic Typology* 6: 49–93.
—— 2003. *The Paradigmatic Structure of Person Marking*. Oxford: Oxford University Press.

D

Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.
—— 1990. 'Standard Average European as an Exotic Language'. In Bechert, Johannes, Giuliano Bernini, and Claude Buridant, eds., *Toward a Typology of European Languages*, Empirical Approaches to Language Typology 8, pp. 3–8. Berlin: Mouton de Gruyter.
—— 1995. 'Areal Tendencies in Tense-Aspect Systems'. In Valentina Bianchi, Pier Marco Bertinetto, Östen Dahl, and Mario Squartini, eds., *Temporal Reference, Aspect and Actionality*, pp. 11–28. Torino: Rosenberg & Sellier.
—— 2000. 'The Tense-Aspect Systems of European Languages in a Typological Perspective'. In Östen Dahl, ed., *Tense and Aspect in the Languages of Europe*, pp. 3–25. Berlin: Mouton de Gruyter.
Daniel, Mixail. 2000. 'Tipologija associativnoj množestvennosti. [The typology of associative plurals]'. Russian State University for Humanities Doctoral Dissertation.
Daniels, Peter T. and William Bright, eds., 1996. *The World's Writing Systems*. New York: Oxford University Press.
d'Ans, André-Marcel. 1968. *Le créole français d'Haïti: Étude des unités d'articulation, d'expansion et de communication*. Janua Linguarum, Series Practica 106. The Hague and Paris: Mouton.

Darwin, Charles. 1872. *The Expression of the Emotions in Man and Animals*. Reprinted by the University of Chicago Press, Chicago and London, 1965.

Das Gupta, Kamalesh. 1971. *An Introduction to the Nocte Language*. Shillong: North-East Frontier Agency.

Dasgupta, Dipankar and Suhnu Ram Sharma. 1982. *Handbook of the Önge Language*. Linguistic Series. Anthropological Survey of India 5. Calcutta: Anthropological Survey of India.

Davies, John. 1981. *Kobon*. Lingua Descriptive Studies 3. Amsterdam: North-Holland.

—— 1989. *Kobon*. London: Routledge.

Davies, William. 1984. 'Antipassive: Choctaw Evidence for a Universal Characterization'. In David Perlmutter and Carol Rosen, eds., *Studies in Relational Grammar 2*, pp. 331–76. Chicago: University of Chicago Press.

—— 1999. *Madurese*. München and Newcastle: Lincom Europa.

Davis, Karen. 2003. *A Grammar of the Hoava Language, Western Solomons*. Pacific Linguistics, 535. Canberra: Australian National University.

Davison, Alice. 2000. 'Lexicon Anaphora and Pronouns in Hindi/Urdu'. In B. C. Lust, Kashi Wali, J. W. Gair, and K. V. Subbarao, eds., *Lexical Anaphors and Pronouns in Selected South Asian Languages: a Principled Typology*, Empirical Approaches to Language Typology 22, pp. 431–2. Berlin and New York: Mouton de Gruyter.

Day, Christopher. 1973. *The Jacaltec Language*. Indiana University Publications, Language Science Monographs 12. Bloomington: Indiana University.

Dayley, Jon P. 1985. *Tzutujil Grammar*. University of California Publications in Linguistics 107. Berkeley and Los Angeles: University of California Press.

—— 1989a. *Tümpisa (Panamint) Shoshone Dictionary*. Berkeley: University of California Press.

—— 1989b. *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.

de Haan, Ferdinand. 1999. Evidentiality and Epistemic Modality: Setting Boundaries. *Southwest Journal of Linguistics* 18: 83–101.

—— 2001. 'The Cognitive Basis of Visual Evidentials'. In Alan Cienki, Barbara J. Luka, and Michael B. Smith, eds., *Conceptual and Discourse Factors in Linguistic Structure*, pp. 91–106. Stanford: CSLI Publications.

de Vries, Lourens. 1989. *Studies in Wambon and Kombai: Aspects of Two Papuan Languages of Irian Jaya*. Amsterdam: Faculty of Arts, University of Amsterdam.

—— 1993. *Forms and Functions in Kombai, an Awyu Language of Irian Jaya*. Pacific Linguistics, Series B-108. Canberra: Australian National University.

Décsy, Gyula. 1966. *Yurak Chrestomathy*. Bloomington: Indiana University Press.

Dedrick, John M., and Eugene H. Casad. 1999. *Sonora Yaqui Language Structures*. Tucson: University of Arizona Press.

Delbrück, Berthold. 1893. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen. Volume 3: Syntax der indogermanischen Sprachen 1*. Strassburg: Trübner.

Dench, Alan Charles. 1991. 'Panyjima'. In R. M. W. Dixon and Barry J. Blake, eds., *Handbook of Australian Languages 4*, pp. 124–243. Melbourne: Oxford University Press.

Derbyshire, Desmond C. 1979. *Hixkaryana*. Lingua Descriptive Studies 1. Amsterdam: North-Holland Publishing Company.

—— 1985. *Hixkaryana and Linguistic Typology*. Dallas: Summer Institute of Linguistics.

Derbyshire, William W. 1993. *A Basic Reference Grammar of Slovene*. Columbus, Ohio: Slavica.

Deshpande, Madhav M. 1992. 'Justification for Verb-Root Suppletion in Sanskrit'. *Historische Sprachforschung* 105: 18–49.

Deza Galindo, Juan Francisco. 1992. *Gramática de la lengua aymara*. Lima: Artex Editores.

Diagana, Ousmane Moussa. 1995. *La langue soninkée. Morphosyntaxe et sens à travers le parler de Kaédi (Mauritanie)*. Paris: L'Harmattan.

Dickens, Patrick. 1992. *Ju|'hoan Grammar*. Windhoek: Nyae Development Foundation.

—— n. d. *Ju|'hoan Grammar*. Unpublished Manuscript.

Diessel, Holger. 1999. *Demonstratives: Form, Function, and Grammaticalization*. Typological Studies in Language 42. Amsterdam: John Benjamins.

—— 2001. 'The Ordering Distribution of Main and Subordinate Clauses: a Typological Study'. *Language* 77: 433–55.

Dik, Simon C. 1989. *The Theory of Functional Grammar*. Dordrecht: Foris.

—— 1997a. *The Theory of Functional Grammar. Volume 1: The Structure of the Clause*. Berlin and New York: Mouton de Gruyter.

—— 1997b. *The Theory of Functional Grammar. Volume 2: Complex and Derived Constructions*. Berlin and New York: Mouton de Gruyter.

Dillon, Myles and Donncha Ó Cróinín. 1961. *Irish*. Teach Yourself Books. London: English Universities Press.

Dimmendaal, G. J. 1983. *The Turkana Language*. Dordrecht: Foris Publications Holland.

Ding, Sizhi. 1998. 'Fundamentals of Prinmi (Pumi): A Tibeto-Burman Language of Northwestern Yunnan, China'. Australian National University Ph.D. Thesis.

Dixon, R. M. W. 1977. *A Grammar of Yidiɲ*. Cambridge Studies in Linguistics 19. Cambridge: Cambridge University Press.

—— 1979. 'Ergativity'. *Language* 55: 59–138.

—— 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

—— 1982. *'Where Have All the Adjectives Gone?' and other Essays in Semantics and Syntax*. Berlin: Mouton de Gruyter.

—— 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

—— 1994. *Ergativity*. Cambridge: Cambridge University Press.

—— 2000. 'A Typology of Causatives: Form, Syntax and Meaning'. In R. M. W. Dixon and Alexandra Y. Aikhenvald, eds., *Changing Valency: Case Studies in Transitivity*, pp. 30–83. Cambridge: Cambridge University Press.

—— and Alexandra Y. Aikhenvald. 1999. *The Amazonian Languages*. Cambridge: Cambridge University Press.

—— —— 2002. *Word: A Cross-linguistic Typology*. Cambridge: Cambridge University Press.

Dobrin, Lise. 2001. Arapesh. In Jane Garry and Carl Rubino, eds., *Encyclopedia of the World's Languages: Past and Present*, pp. 33–8. New York and Dublin: H. W. Wilson Press.

Dobrušina, N. R. 2001. 'Formy imperativnoj serii'. In A. E. Kibrik, ed., *Bagvalinskij jazyk: grammatika, teksty, slovari*, pp. 319–32. Moscow: Nasledie.

Doke, Clement M. 1953. *Zulu-English Dictionary*. Johannesburg: Witwatersrand University Press.

Dol, Philomena. 1999. 'A Grammar of Maybrat: A Language of the Bird's Head, Irian Jaya, Indonesia'. University of Leiden Ph.D. Dissertation.

Donaldson, Tamsin. 1980. *Ngiyambaa: the Language of the Wangaaybuwan*. Cambridge Studies in Linguistics 29. Cambridge: Cambridge University Press.

Donohue, Mark. 1999. *A Grammar of Tukang Besi*. Berlin and New York: Mouton de Gruyter.

Dorais, Louis-Jacques. 1988. *Tukilik: an Inuktitut Grammar for All*. Quebec: Association Inuksiutiit Katimajiit Inc., Groupes d'études inuit et circumpolaires, Université Laval.

Dougherty, Janet W. D. 1983. *West Futuna-Aniwa: An Introduction to a Polynesian Outlier Language*. University of California Publications in Linguistics 102. Berkeley: University of California Press.

Douglas, Bryan. 1991. 'Aspects of Tugun Phonology and Syntax'. University of Texas at Arlington M.A. Thesis.

Douglas, W. H. 1959. *An Introduction to the Western Desert Language*. Sydney: University of Sydney.

Dowty, David. 1991. 'Thematic Proto-roles and Argument Selection'. *Language* 67: 547–619.

Drabbe, Peter. 1952. *Spraakkunst van het Ekagi*. 'S-Gravenhage: Martinus Nijhoff.

Drake, John. 1903. *A Grammar of the Kurku Language*. Calcutta: Baptist Mission Press.

Dressler, Wolfgang. 1985. 'Suppletion in Word-Fomation'. In J. Fisiak, ed., *Historical Morphology*, pp. 97–112. The Hague: Mouton.

—— Willi Mayerthaler, Oswald Panagl, and Wolfgang U. Wurzel. 1987. *Leitmotifs in Natural Morphology*. Amsterdam: John Benjamins.

Dryer, Matthew S. 1983. 'Indirect Objects in Kinyarwanda Revisited'. In David M. Perlmutter, ed., *Studies in Relational Grammar 1*, pp. 129–40. Chicago: University of Chicago Press

—— 1986. 'Primary Objects, Secondary Objects, and Antidative'. *Language* 62: 808–45.

—— 1988a. 'Object-Verb Order and Adjective-Noun Order: Dispelling a Myth'. *Lingua* 74: 77–109.

—— 1988b. 'Universals of Negative Position'. In Michael Hammond, Edith Moravcsik, and Jessica Wirth, eds., *Studies in Syntactic Typology*, pp. 93–124. Amsterdam: John Benjamins.

—— 1989a. 'Large linguistic areas and language sampling'. *Studies in Language* 13: 257–92.

—— 1989b. 'Salinan Word Order'. In Scott Delancey, ed., *Papers from the 1988 Hokan-Penutian Languages Workshop*. University of Oregon Papers in Linguistics, Publications of the Center for Amerindian

Linguistics and Ethnography 1, pp. 40–49. Eugene, Oregon: Department of Linguistics, University of Oregon.
—— 1991. 'SVO Languages and the OV:VO Typology'. *Journal of Linguistics* 27: 443–82.
—— 1992. 'The Greenbergian Word Order Correlations'. *Language* 68: 81–138.
—— 1997. 'On the Six-Way Word Order Typology'. *Studies in Language* 21: 69–103.
Du Feu, Veronica. 1996. *Rapanui*. Descriptive Grammar Series. London: Routledge.
Dugast, Idelette. 1971. *Grammaire du tunen*. Langues et Littératures de l'Afrique Noire VIII. Paris: Edition Klincksieck.
Dumézil, Georges. 1931. *La langue des Oubykhs*. Paris: Librairie ancienne Honoré Champion.
Dunn, Michael. 1999. 'A Grammar of Chukchi'. Australian National University Ph.D. Thesis.
—— 2000. 'Chukchi Women's Language: A Historical-Comparative Perspective'. *Anthropological Linguistics* 42: 305–28.
Durie, Mark. 1986. 'The Grammaticization of Number as a Verbal Category'. In Vassiliki Nikiforidou, Mary VanClay, Mary Niepokuj, and Deborah Feder, eds., *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society, 15–17 February 1986*, pp. 355–70. Berkeley, California: Berkeley Linguistics Society Publications.
Durkheim, Émile. 1915. *The Elementary Forms of the Religious Life*. London: G. Allen & Unwin.

E

Eades, Diana. 1979. 'Gumbainggir'. In R. M. W. Dixon and B. J. Blake, eds., *Handbook of Australian Languages 1*, pp. 244–361. Amsterdam: John Benjamins.
Ebermann, Erwin. 1986. 'Die Sprache der Mauka'. Dissertationen der Universität Wien. Wien: Verband der wissenschaftlichen Gesellschaft Österreichs (VWGÖ).
Ebert, Karen. 2001. 'Tense-Aspect Flip-Flop and a somewhat Elusive Gram Type'. In Karen H. Ebert and Fernando Zúñiga, eds., *Aktionsart and Aspectotemporality in Non-European Languages*, pp. 141–58. Zürich: Universität Zürich, Seminar für Allgemeine Sprachwissenschaft.
Echols, John M. and Hassan Shadily. 1961/1989. *An Indonesian-English Dictionary*. Jakarta: Gramedia. [3rd edn.]
Edmonson, Barbara Wedemeyer. 1988. 'A Descriptive Grammar of Huastec (Potosino Dialect)'. Tulane University Dissertation.
Egli, Hans. 1990. *Paiwangrammatik*. Wiesbaden: Otto Harrassowitz.
Ehrman, Madeline Elizabeth. 1970. *Contemporary Cambodian: Grammatical Sketch*. Washington, DC: Foreign Service Institute.
Elbert, Samuel H. 1974. *Puluwat Grammar*. Pacific Linguistics, Series B-9. Canberra: Australian National University.
—— and Mary Kawena Pukui. 1979. *Hawaiian Grammar*. Honolulu: University Press of Hawaii.
Ellis, Clarence D. 1983. *Spoken Cree*. West Coast of James Bay: Pica Pica Press. [2nd edn.]
Elwell-Sutton, L. P. 1941. *Colloquial Persian*. London: Routledge and Kegan Paul Limited.
Emeneau, Murray B. 1980. 'The Indian Linguistic Area Revisited'. In M. B. Emeneau, ed., *Language and Linguistic Area*, pp. 197–249. Stanford: Stanford University Press.
Engel, Ralph, and Robert E. Longacre. 1963. 'Syntactic Matrices in Ostuacan Zoque'. *International Journal of American Linguistics* 29: 331–44.
England, Nora C. 1983a. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.
—— 1983b. 'Ergativity in Mamean (Mayan) Languages'. *International Journal of American Linguistics* 49: 1–19.
—— 1988. Mam Voice. In Masayoshi Shibatani, ed., *Passive and Voice*, pp. 525–45. Amsterdam: John Benjamins.
Erwin, Wallace M. 1963. *A Short Reference Grammar of Iraqi Arabic*. Washington, DC: Georgetown University Press.
Evans, Nicholas D. 1995. *A Grammar of Kayardild: with Historical-Comparative Notes on Tangkic*. Berlin and New York: Mouton de Gruyter.
—— 2003. 'Comparative non-Pama-Nyungan and Australian historical linguistics'. In Nicholas Evans (ed.), *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*, pp. 3–25. Canberra: Australian National University.
—— Dunstan Brown and Greville G. Corbett. 2002. 'The Semantics of Gender in Mayali: Partially Parallel Systems and Formal Implementation'. *Language* 77: 111–55.
Everaert, Martin. 1999. 'Types of anaphoric expressions: reflexives and reciprocals'. In Zygmunt Frajzyngier and Traci S. Curl, eds., *Reciprocals: Forms and Functions*, Typological Studies in Language 41, pp. 63–84. Amsterdam: John Benjamins.
Everett, Daniel L. 1986. 'Pirahã'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 1*, pp. 200–325. Berlin: Mouton de Gruyter.
—— and Barbara Kern. 1997. *Wari': the Pacaas Novos Language of Western Brazil*. Descriptive Grammar Series. London: Routledge.
Ezard, Bryan. 1997. *A Grammar of Tawala, an Austronesian Language of the Milne Bay Area, Papua New Guinea*. Pacific Linguistics, Series C-137. Canberra: Australian National University.

F

Facundes, Sidney da Silva. 2000. 'The Language of the Apurinã People of Brazil (Maipure/Arawak)'. SUNY at Buffalo Ph.D. Dissertation.
Faltz, Leonard M. 1985. *Reflexivization: A Study in Universal Syntax*. New York: Garland.
Faraclas, N. 1984. *A Grammar of Obolo*. Bloomington: Indiana University Linguistics Club.
Feev, Boris B. 1998. *Akustičeskie xarakteristiki glasnyx ketskogo jazyka (Pakulixinskij govor)*. Novosibirsk: Rossijskaja Akademija Nauk.
Feldman, Harry. 1986. *A Grammar of Awtuw*. Pacific Linguistics, Series B-94. Canberra: Australian National University.
Ferguson, Charles A. 1966. 'Assumptions about Nasals: a Sample Study in Phonological Universals'. In Joseph H. Greenberg, ed., *Universals of Language*, pp. 53–60. Cambridge, MA: MIT Press. [2nd edn.]
—— Larry M. Hyman, and John Ohala, eds. 1975. *Nasàlfest*. Stanford: Language Universals Project, Department of Linguistics, Stanford University.
Fernandez, Frank. 1967. 'A Grammatical Sketch of Remo: A Munda Language'. University of North Carolina at Chapel Hill Ph.D. Dissertation.
Fertig, David. 1998. 'Suppletion, Natural Morphology, Diagrammaticity'. *Linguistics* 36.6: 1065–91.
Fife, James. 1990. *The Semantics of the Welsh Verb. A Cognitive Approach*. Cardiff: University of Wales Press.
Fleisch, Axel. 2001. *Lucazi Grammar. A Morphosemantic Analysis*. Köln: Rüdiger Köppe.
Floyd, Rick. 1999. *The Structure of Evidential Categories in Wanka Quechua*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
Foley, William A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
—— 1991. *The Yimas Language of Papua New Guinea*. Stanford: Stanford University Press.
—— 2000. 'The languages of New Guinea'. *Annual Review of Anthropology* 29: 357–404.
—— and Robert D. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
Forchheimer, Paul. 1953. *Category of Person in Language*. Berlin: de Gruyter.
Foreman, Velma M. 1974. *Grammar of Yessan-Mayo*. Language Data, Asian-Pacific Series 4. Santa Ana: Summer Institute of Linguistics.
Fortescue, Michael. 1984. *West Greenlandic*. Croom Helm Descriptive Grammars. London: Croom Helm.
—— (forthcoming.) 'Reciprocals in West Greenlandic Eskimo'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Fortune, Reo F. 1942. *Arapesh*. Publications of the American Ethnological Society 19. New York: J. J. Augustin. [Reprinted 1977, New York: AMS Press.]
Frachtenberg, Leo J. 1913. *Coos Texts*. Columbia University Contributions to Anthropology. New York: Columbia University Press.
—— 1922a. 'Siuslawan (Lower Umpqua)'. In Franz Boas, ed., *Handbook of American Indian Languages 2*, Smithsonian Institution Bureau of American Ethnology Bulletin 40, pp. 431–629. Washington: Government Printing Office.
—— 1922b. 'Coos'. In Franz Boas, ed., *Handbook of American Indian Languages 2*, Smithsonian Institution Bureau of American Ethnology Bulletin 40, pp. 297–429. Washington: Government Printing Office.
Frajzyngier, Zygmunt. 1989. *A Grammar of Pero*. Berlin: Dietrich Reimer Verlag.
—— 1993. *A Grammar of Mupun*. Berlin: Dietrich Reimer Verlag.
Frank, Paul. 1990. *Ika Syntax*. Studies in the Language of Colombia 1. Arlington: Summer Institute of Linguistics and the University of Texas at Arlington.
Frank, Wright. 1999. 'Nuer Noun Morphology'. University at Buffalo M.A. Thesis.

Franklin, Karl James. 1971. *A Grammar of Kewa, New Guinea*. Pacific Linguistics, Series C-16. Canberra: Australian National University.
Frantz, Donald. 1991. *Blackfoot Grammar*. Toronto: University of Toronto Press.
Fraser, Norman M. and Greville G. Corbett. 1995. 'Gender, Animacy and Declensional Class Assignment: a Unified Account for Russian'. In Geert Booij and Jaap van Marle, eds., *Yearbook of Morphology 1994*, pp. 123–50. Dordrecht: Kluwer.
—— —— 1997. 'Defaults in Arapesh'. *Lingua* 103: 25–57.
Friberg, Barbara. 1996. 'Konjo's Peripatetic Person Markers'. In H. Steinhauer, ed., *Papers in Austronesian Linguistics 3*, Pacific Linguistics, Series A-84, pp. 137–71. Canberra: Research School of Pacific and Asian Studies, Australian National University.
Friedrich, Paul. 1966. *Structural Implications of Russian Pronominal Usage*. Janua Linguarum XX. The Hague and Paris: Mouton.
—— 1972. 'Social Context and Semantic Features: the Russian Pronominal Usage'. In J. J. Gumperz and D. Hymes, eds., *Directions in Sociolinguistics*, pp. 270–300. New York: Holt Rinehardt and Winston.
Frishberg, Nancy. 1987. 'Ghanaian Sign Language'. In John van Cleve, ed., *Gallaudet Encyclopedia of Deaf People and Deafness* 3, pp. 778–9. New York: McGraw-Hill.
Fromm, Hans. 1982. *Finnische Grammatik*. Heidelberg: Winter.
—— and Matti Sadeniemi. 1956. *Finnisches Elementarbuch, Volume 1: Grammatik*. Heidelberg: Winter.

G

Gabas, Nilson, Jr. 1999. 'A Grammar of Karo, Tupi (Brazil)'. University of California at Santa Barbara Ph.D. Dissertation.
Gadžiaxmedov, N. E. 2000. *Slovoizmenitel'nye kategorii imeni i glagola v kumykskom jazyke*. Maxačkala.
Gair, James W. 1970. *Colloquial Sinhalese Clause Structures*. The Hague: Mouton.
Gamble, Geoffrey. 1978. *Wikchamni Grammar*. University of California Publications in Linguistics 89. Berkeley and Los Angeles: University of California Press.
Garry, Jane and Carl Rubino, eds. 2001. *Encyclopedia of the World's Languages: Past and Present*. New York and Dublin: H. W. Wilson Press.
Gary, Judith Olmsted and Saad Gamal-Eldin. 1982. *Cairene Egyptian Colloquial Arabic*. Lingua Descriptive Studies 6. Amsterdam: North Holland.
—— and Edward L. Keenan. 1977. 'On Collapsing Grammatical Relations in Universal Grammar'. In Peter Cole and Jerrold Sadock, eds., *Syntax and Semantics, Volume 8: Grammatical Relations*, pp. 83–120. New York: Academic Press.
Geniušienė, Emma. 1987. *The Typology of Reflexives*. Empirical Approaches to Language Typology 2. Berlin: Mouton de Gruyter.
—— (forthcoming). 'Reciprocal and Reflexive Constructions in Lithuanian (with references to Latvian)'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Gensler, Orin D. 1994. 'On Reconstructing the Syntagm S-Aux-O-V-Other to Proto-Niger-Congo'. *Proceedings of the Berkeley Linguistics Society* 20: 1–20. (Special Session on Historical Issues in African Linguistics).
—— 1997. 'Grammaticalization, Typology, and Niger-Congo Word Order: Progress on a still-unsolved Problem. Review article on *Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen*, by Ulrike Claudi'. *Journal of African Languages and Linguistics* 18: 57–93.
Gerdts, Donna. 1988. *Object and Absolutive in Halkomelem Salish*. New York: Garland.
—— 1993. Mapping Halkomelem Grammatical Relations. *Linguistics* 31: 591–622.
—— and Thomas Hukari. (forthcoming). *Multiple Antipassives in Halkomelem Salish*. Berkeley Linguistic Society. Proceedings of the 26th Special Session. University of California, Berkeley.
Gibson, Jeanne D. 1980. 'Clause Union in Chamorro and Universal Grammar'. University of California San Diego Ph.D. Dissertation.
Gil, David. 1982. 'Distributive Numerals'. University of California, Los Angeles Ph.D. Dissertation.
—— 1987. 'Definiteness, Noun-Phrase Configurationality, and the Count-Mass Distinction'. In E. J. Reuland and A. G. B. ter Meulen, eds., *The Representation of (In)definiteness*, pp. 254–69. Cambridge, MA: MIT Press.
—— 1988. 'Georgian Reduplication and the Domain of Distributivity'. *Linguistics* 26: 1039–65.
—— 1989. 'Korean *-ssik*. A Syntactic and Semantic Study'. In S. Kuno, I.-H. Lee, J. Whitman, S.-Y. Bak, Y.-S. Kang and Y.-J. Kim, eds., *Harvard Studies in Korean Linguistics 3*, pp. 423–30. Seoul: Hanshin Publishing Company.
—— 1992. 'Scopal Quantifiers: Some Universals of Lexical Effability'. In M. Kefer and J. van der Auwera, eds., *Meaning and Grammar, Cross-Linguistic Perspectives*, pp. 303–45. Berlin: Mouton de Gruyter.
—— 1993. 'Nominal and Verbal Quantification'. *Sprachtypologie und Universalienforschung* 46: 275–317.
—— 1994a. 'Conjunctive Operators: A Unified Semantic Analysis'. In Peter Bosch and Rob van der Sandt, eds., *Focus and Natural Language Processing, Volume 2: Semantics*, pp. 311–22. Proceedings of a Conference in Celebration of the 10th Anniversary of the *Journal of Semantics*, 12th–15th June, 1994, Hotel Schloß Wolfsbrunnen, Meinhard-Schwebda, Federal Republic of Germany, Working Papers of the Institute for Logic and Linguistics, Working Paper 7. Heidelberg.
—— 1994b. 'Conjunctive Operators in South-Asian Languages'. In A. Davison and F. M. Smith, eds., *Papers from the Fifteenth South Asian Language Analysis Round Table Conference*, South Asian Studies Program, pp. 82–105. Iowa City: University of Iowa.
—— 1994c. 'Everything Also Must Grab: A Unified Semantic Analysis of Singlish "Also" '. In F. Plank, ed., *Agreement Gender Number Genitive*, EUROTYP Working Papers, Series 7, Number 23, pp. 133–45. Berlin: European Science Foundation, EUROTYP Programme.
—— 1994d. 'Adjectives without Nouns: An E-mail Conversation'. In F. Plank, ed., *Conversations on Noun Phrases*, EUROTYP Working Papers, Series 7, Number 20, pp. 31–48. Berlin: European Science Foundation EUROTYP Programme.
—— 1996. 'Maltese "Collective Nouns": A Typological Perspective'. *Rivista di Linguistica* 8: 53–87.
—— 2001. 'Escaping Eurocentrism: Fieldwork as a Process of Unlearning'. In P. Newman and M. Ratliff, eds., *Linguistic Fieldwork*, pp. 102–32. Cambridge: Cambridge University Press.
Ginter, Károly and László Tarnói. 1991. *Ungarisch für Ausländer*. Budapest: Tankönyvkiadó.
Giridhar, Puttushetra Puttuswamy. 1980. *Angami Grammar*. Manasagangotri, Mysore: Centeral Institute of Indian Languages.
—— 1994. *Mao Naga Grammar*. Mysore: Central Institute of Indian Languages.
Givón, Talmy. 1978. 'Negation in Language: Pragmatics, Function, Ontology'. In Peter Cole, ed., *Syntax and Semantics, Volume 9: Pragmatics*, pp. 69–112. New York: Academic Press.
—— 1980. 'The Binding Hierarchy and the Typology of Complements'. *Studies in Language* 4: 333–77.
—— 1981. 'On the Development of the Numeral "One" as an Indefinite Marker'. *Folia Linguistica Historica* 2: 35–53.
Givón, Talmy, ed. 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.
Givón, Talmy. 1984. *Syntax: a Functional-Typological Introduction 1*. Amsterdam: John Benjamins.
—— 1990. *Syntax: a Functional-Typological Introduction 2*. Amsterdam: John Benjamins.
—— 1994. *Voice and Inversion*. Amsterdam: John Benjamins.
—— 2001. *Syntax: An Introduction 1*. Amsterdam: John Benjamins.
—— (in collaboration with the Southern Ute Tribe). 1980. *Ute Reference Grammar*. Ignacio, Colorado: Ute Press, The Southern Ute Tribe.
Gnerre, Maurizio. 1999. *Profilo Descrittivo e Storico-comparativo di una Lingua Amazzonica: lo Shuar (Jívaro)*. Cuaderni di AION, nuova serie 1. Naples: Oriental Institute.
Goddard, Cliff. 1982. 'Case Systems and Case Marking in Australian Languages: a New Interpretation'. *Australian Journal of Linguistics* 2: 167–96.
—— 1983. 'A Semantically-oriented Grammar of the Yankunytjara Dialect of the Western Desert Language'. Australian National University Ph.D. Thesis.
Goddard, Ives. 1996. 'Native languages and language families of North America'. Map accompanying Ives Goddard, ed., *Handbook of North American Indians*, Volume 17: Languages. Washington: Smithsonian Institution.
Goedemans, R. W. N. 1996. 'Syllable Weight and Prominence'. In R. W. N. Goedemans, H. G. van der Hulst, and E. A. M. Visch, eds., *Stress Patterns of the World, Volume 1: Background*, HIL Publications 2, pp. 115–64. The Hague: Holland Academic Graphics.
—— 1997. 'Putting the Mathimathi Stress Rule in its Proper Perspective'. *The Australian Journal of Linguistics* 17: 43–69.
—— 1998. *Weightless Segments: A Phonetic and Phonological Study Concerning the Metrical Irrelevance of Syllable Onsets*. LOT Dissertations 9. The Hague: Holland Academic Graphics.
—— and H. van der Hulst. (forthcoming.) *Stress Patterns of the World: The Data*.

Gordon, Lynn. 1986. *Maricopa Morphology and Syntax*. University of California Publications in Linguistics 108. Berkeley: University of California Press.
Gowda, K. S. G. 1975. *Ao Grammar*. Manasagangotri, Mysore: Central Institute of Indian Languages.
Green, M. M. and G. E. Igwe. 1963. *A Descriptive Grammar of Igbo*. London: Oxford University Press.
Green, Rebecca. 1995. 'A Grammar of Gurr-goni (North Central Arnhem Land)'. Australian National University Ph.D. Thesis.
Greenberg, Joseph H. 1957. 'Order of Affixing: a Study in General Linguistics'. In Joseph H. Greenberg, ed., *Essays in Linguistics*, pp. 86–94. Chicago: University of Chicago Press.
—— 1963. 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements'. In Joseph H. Greenberg, ed., *Universals of Human Language*, pp. 73–113. Cambridge, MA: MIT Press.
—— 1978. 'Generalizations about Numeral Systems'. In Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, eds., *Universals of Human Language, Volume 3: Word Structure*, pp. 249–95. Stanford: Stanford University Press.
Greene, David. 1966. *The Irish Language*. Dublin: The Three Candles.
Gregores, Emma and Jorge A. Suárez. 1967. *A Description of Colloquial Guaraní*. The Hague: Mouton.
Grimes, Barbara F. (ed.) 2000. *Ethnologue*. 2 vols. 14th edition. Dallas: SIL International.
Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
Grinevald, Colette G. (n. d.). *A Grammar of Rama*. Report of National Science Foundation BNS 8511156. Unpublished Draft Manuscript, University of Lyon.
Groote Eylandt Linguistics-langwa. 1993. *Eningerribirra-langwa jurra*. Angurugu: Groote Eylandt Linguistics.
Groves, Terab'ata R., Gordon W. Groves, and Roderick Jacobs. 1985. *Kiribatese: An Outline Description*. Pacific Linguistics, Series D-64. Canberra: Australian National University.
Gruzdeva, Ekaterina. 1998. *Nivkh*. Languages of the World/Materials 111. München and Newcastle: Lincom Europa.
Guentchéva, Zlatka (ed.). 1996. *L'énonciation médiatisée*. Louvain: Peeters.
Guentchéva, Zlatka and Nicole Rivière. (forthcoming.) 'Reciprocal and Reflexive Constructions in French'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Guérin, Françoise. 1998. 'Première approche du système verbal de l'ingouche'. In Fernand Bentolila, ed., *Systèmes verbaux*, pp. 253–72. Louvain-la-Neuve: Peeters.
Guirardello, Raquel. 1999. 'A Reference Grammar of Trumai'. Rice University Ph.D. Thesis.
Güldemann, Tom. 2000. 'Noun Categorization in Non-Khoe Lineages of Khoisan'. *Afrikanistische Arbeitspapiere* (Köln) 63: 5–33.
Guma, S. M. 1971. *An Outline Structure of Southern Sotho*. Pietermaritzburg: Shuter and Shooter.

H

Haarmann, Harald. 1970. *Die indirekte Erlebnisform als grammatische Kategorie. Eine eurasische Isoglosse*. Wiesbaden: Harrassowitz.
Haas, Mary R. 1940. *Tunica*. New York: J. J. Augustin Publisher.
—— 1942. 'The Use of Numeral Classifiers in Thai'. *Language* 18: 201–6.
Hagman, Roy Stephen. 1977. *Nama Hottentot Grammar*. Research Center for Language and Semiotic Studies. Bloomington: Indiana University Press.
Haiman, John. 1983. 'Iconic and Economic Motivation'. *Language* 59: 781–819.
—— 1985. *Natural Syntax*. Cambridge: Cambridge University Press.
Hajek, John. 1997. *Universals of Sound Change in Nasalisation*. Oxford: Blackwell.
—— and R. W. N. Goedemans. 2003. 'Word-initial Geminates and Stress in Pattani Malay'. *The Linguistic Review* 20.1: 79–94.
Hale, Kenneth. 1983. 'Walpiri and the Grammar of Nonconfigurational Languages'. *Natural Language and Linguistic Theory* 1: 5–47.
Hall, Christopher. 1988. 'Integrating Diachronic and Processing Principles in Explaining the Suffixing Preference'. In John A. Hawkins, ed., *Explaining Language Universals*, pp. 321–49. Oxford: Basic Blackwell.
Haller, Felix. 2000. *Dialekt und Erzählungen von Shigatse*. Bonn: VGH Wissenschaftsverlag.
Hamel, Patricia J. 1994. *A Grammar and Lexicon of Loniu, Papua New Guinea*. Pacific Linguistics, Series C-103. Canberra: Australian National University.
Hamilton, R. C. 1900. *An Outline Grammar of the Dafla Language*. Shillong: The Assam Secretariat Printing Office.
Hanoteau, Adolphe. 1896. *Essai de grammaire de la langue tamachek'*. Algiers: Jourdan.
—— 1906. *Essai de grammaire kabyle*. Algiers: Jourdan.
Hardin, C. L. and Luisa Maffi, eds. 1997. *Color Categories in Thought and Language*. Cambridge: Cambridge University Press.
Hardman, Martha J. 2000. *Jaqaru*. Languages of the World/Materials 183. München: Lincom Europa.
Hardman-de-Bautista, Martha J. 1974. *Outline of Aymara, Phonological and Grammatical Structure*. Gainesville: University of Florida.
Hardy, Heather K. and Timothy Montler. 1988. 'Imperfective Gemination in Alabama'. *International Journal of American Linguistics* 54: 399–475.
Harms, Philip Lee. 1994. *Epena Pedee Syntax*. Studies in the Languages of Colombia 4. SIL and University of Texas Publications in Linguistics 118. Dallas and Arlington: Summer Institute of Linguistics and University of Texas.
Harrell, Richard S. 1962. *A Short Reference Grammar of Moroccan Arabic*. Washington: Georgetown University Press.
Harriehausen, Bettina. 1990. *Hmong Njua: Syntaktische Analyse einer gesprochenen Sprache mithilfe daten-verarbeitungstechnischer Mittel und sprachvergleichende Beschreibung des südostasiatischen Sprachraumes*. Tübingen: Max Niemeyer.
Harris, Alice. 1981. *Georgian Syntax*. Cambridge: Cambridge University Press.
Harris, Martin. 1988. 'French'. In Martin Harris and Nigel Vincent, eds., *The Romance Languages*, pp. 209–45. London: Croom Helm.
Harrison, Carl H. 1986. 'Verb Prominence, Verb Initialness, Ergativity and Typological Disharmony in Guajajara'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 1*, pp. 407–39. Berlin: Mouton de Gruyter.
Harrison, David. 2000. *Topics in the Phonology and Morphology of Tuvan*. Yale University Ph.D. Dissertation.
Harrison, Roy, Margaret Harrison, and H. Castulo Garcia. 1981. *Diccionario Zoque de Copainalá*. Serie de vocabularios y diccionarios indígenas "Mariano Silva y Aceves" 23. Mexico: Instituto Lingüístico de Verano.
Harrison, Sheldon P. 1973. 'Reduplication in Micronesian Languages'. *Oceanic Linguistics* 12.1/2: 407–54.
—— with the assistance of Salich Y. Albert. 1976. *Mokilese Reference Grammar*. Honolulu: University Press of Hawaii.
Hartzler, Margaret. 1994. 'Sentani'. In Peter Kahrel and René van den Berg, eds., *Typological Studies in Negation*, Typological Studies in Language 29, pp. 51–64. Amsterdam: John Benjamins.
Harvey, Mark. 1992. 'The Gaagudju People and their Language'. University of Sydney Ph.D. Dissertation.
Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Mouton Grammar Library 9. Berlin: Mouton de Gruyter.
—— 1995. 'The Converb as a Cross-Linguistically Valid Category'. In M. Haspelmath and E. König, eds., *Converbs in Cross-Linguistic Perspective*, pp. 1–55. Berlin: Mouton de Gruyter.
—— 1996. 'Word-class-changing Inflection and Morphological Theory'. In Geert Booij and Jaap van Marle, eds., *Yearbook of Morphology 1995*, pp. 43–66. Dordrecht: Kluwer.
—— 1997. *Indefinite Pronouns*. Oxford: Oxford University Press.
—— 2001. 'The European Linguistic Area: Standard Average European'. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, eds., *Language Typology and Language Universals*, Handbücher zur Sprach- und Kommunikationswissenschaft, pp. 1492–510. Berlin: Mouton de Gruyter.
—— 2002. *Understanding Morphology*. London: Arnold.
—— 2005. 'Coordination'. In Timothy Shopen, ed., *Language Typology and Linguistic Description*. Cambridge: Cambridge University Press. [2nd edn.]
Haugen, Jason. (forthcoming.) 'Reduplicative Allomorphy and Language Prehistory in Uto-Aztecan', In B. Hurch, ed., *Studies in Reduplication*. Berlin: Mouton de Gryuter.
Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.
—— 1990. 'A Parsing Theory of Word Order Universals'. *Linguistic Inquiry* 21: 223–61.
—— 1993. 'Heads, Parsing and Word Order Universals'. In Greville G. Corbett, Norman M. Fraser, and Scott McGlashan, eds., *Heads in Grammatical Theory*, pp. 231–65. Cambridge: Cambridge University Press.
—— 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
—— 2002. 'Symmetries and Asymmetries: their Grammar, Typology, and Parsing'. *Theoretical Linguistics* 28: 95–149.

Hawkins, Robert E. 1998. 'Wai Wai'. In D. C. Derbyshire and G. K. Pullum, eds., *Handbook of Amazonian Languages 4*, pp. 25–224. Berlin: Mouton de Gruyter.
Hayes, Bruce. 1995. *A Metrical Theory of Stress: Principles and Case Studies.* Chicago, Illinois: University of Chicago Press.
Hay-Neave, D. R. 1953. *Lai-Chin Grammar and Exercises*. Rangoon: Govt. Printing and Stationary, Burma.
Hayward, Richard J. 1990a. 'Notes on the Aari Language'. In Richard J. Hayward, ed., *Omotic Language Studies*, pp. 425–93. London: School of Oriental and African Studies, University of London.
—— 1990b. 'Notes on the Zayse Language'. In Richard J. Hayward, ed., *Omotic Language Studies*, pp. 210–355. London: School of Oriental and African Studies, University of London.
He Zili. 1990. '"ni X wo Y": A Formal Idiom in Chinese'. In K. Hall et al., *Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society*, pp. 153–62. Berkeley: Berkeley Linguistics Society.
Head, Brian F. 1978. 'Respect Degrees in Pronominal Reference'. In Joseph Greenberg, ed., *Universals of Human Language, Volume 3: Word Structures*, pp. 151–211. Stanford: Stanford University Press.
Heath, Jeffrey. 1976. 'Antipassivization: A Functional Typology'. *Berkeley Linguistics Society* 2: 202–11.
—— 1978. *Ngandi Grammar, Texts and Dictionary*. Canberra: Australian Institute of Aboriginal Studies.
—— 1980. *Basic Materials in Warndarang: Grammar, Texts, and Dictionary*. Pacific Linguistics, Series B-72. Canberra: Australian National University.
—— 1984. *A Functional Grammar of Nunggubuyu*. Atlantic Highlands NJ: Humanities Press/Canberra: Australian Institute of Aboriginal Studies.
—— 1986. 'Syntactic and Lexical Aspects of Nonconfigurationality in Nunggubuyu (Australia)'. *Natural Language and Linguistic Theory* 4: 295–331.
Heath, Jeffrey, ed. 1998. *Texts in Koyra Chiini: Songhay of Timbuktu, Mali*. Köln: Rüdiger Köppe.
Heath, Jeffrey. 1999a. *A Grammar of Koyraboro (Koroboro) Senni.* Westafrikanische Studien 19. Köln: Rüdiger Köppe Verlag.
—— 1999b. *A Grammar of Koyra Chiini: the Songhay of Timbuktu*. Berlin: Mouton de Gruyter.
Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. New York: Oxford University Press.
—— 1994. 'Areal Influence on Grammaticalization'. In Martin Pütz, ed., *Language Contact and Language Conflict*, pp. 55–68. Amsterdam: John Benjamins.
—— 1997. *Possession: Cognitive Sources, Forces and Grammaticalization*. Cambridge: Cambridge University Press.
—— 1999. 'Polysemy Involving Reflexive and Reciprocal Markers in African Languages'. In Zygmunt Frajzyngier and Traci S. Curl, eds., *Reciprocals: Forms and Functions*, pp. 1–30. Amsterdam and Philadelphia: John Benjamins.
Helmbrecht, Johannes. 2003. 'Politeness Distinctions in Second Person Pronouns'. In Friedrich Lenz, ed., *Deictic Conceptualization of Space, Time and Person*, pp. 185–202. Amsterdam and Philadelphia: John Benjamins.
—— 2004. *Personal Pronouns – Form, Function and Grammaticalization*. University of Erfurt Habilitationsschrift.
Henderson, Eugénie J. A. 1997. *Bwe Karen Dictionary: with Texts and English-Karen Word List*. London: School of Oriental and African Studies, University of London.
Hengeveld, Kees. 1989. 'Layers and Operators in Functional Grammar'. *Journal of Linguistics* 25: 127–57.
—— 1990. 'The Hierarchical Structure of Utterances'. In J. Nuyts, A. M. Bolkestein, and C. Vet, eds., *Layers and Levels of Representation in Language Theory: a Functional View*, pp. 1–24. Amsterdam and Philadelphia: John Benjamins.
—— 1998. 'Adverbial Clauses in the Languages of Europe'. In J. van der Auwera and D. P. O'Baoill, eds., *Adverbial Constructions in the Languages of Europe*, pp. 335–419. Berlin and New York: Mouton de Gruyter.
—— (forthcoming). 'Mood and Modality'. In Geert Booij et al., eds., *Morphology: an International Handbook on Inflection and Word-Formation 2*, Handbooks of Linguistics and Communication Science 17. Berlin and New York: Mouton de Gruyter.
Hercus, Luise A. 1982. *The Bagandji Language*. Pacific Linguistics, Series B-67. Canberra: Australian National University.
—— 1994. *A Grammar of the Arabana-Wangkangurru Language, Lake Eyre Basin, South Australia*. Pacific Linguistics, Series C-128. Canberra: Australian National University.
Hering, Ewald. 1920. *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer. [English version: *Outlines of a Theory of the Light Sense*. 1964. Translated by L. M. Hurvich and Dorothea Jameson. Cambridge, MA: Harvard University Press.]
Heusler, Andreas. 1964. *Altisländisches Elementarbuch*. Heidelberg: Winter.
Hewitt, B. George. 1979. *Abkhaz*. Lingua Descriptive Studies 2. Amsterdam: North Holland Publishing Company.
—— 1987. *The Typology of Subordination in Georgian and Abkhaz*. Berlin and New York: Mouton de Gruyter.
—— 1989. *Abkhaz*. London: Routledge.
—— 1995. *Georgian: A Structural Reference Grammar*. Amsterdam: John Benjamins.
Hill, Deborah. 2002. 'Longgu'. In John Lynch, Malcolm Ross, and Terry Crowley, eds., *The Oceanic Languages*, pp. 538–61. Richmond: Curzon.
Hinds, John. 1986. *Japanese*. Croom Helm Descriptive Grammars. London and Dover, NH: Croom Helm.
Hoa, Monique, Tamara Nikitina, and Vladimir Nedjalkov. (forthcoming.) 'Reciprocal Constructions in Modern Chinese'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Hoàng Thị Chau. 1982. *Grundkurs Vietnamesisch*. Leipzig: Enzyklopädie.
Hoddinott, W. G. and F. M. Kofod. 1988. *The Ngankikurungkurr Language (Daly River, Area, Northern Territory)*. Pacific Linguistics, Series D-77. Canberra: Australian National University.
Hoffmann, J. 1903. *Mundari Grammar*. Calcutta: Bengal Secretariat Press.
Hofling, Charles Andrew. 2000. *Itzaj Maya Grammar*. Salt Lake City: University of Utah Press.
Hoijer, Harry. 1972. *Tonkawa Texts*. University of California Publications in Linguistics 73. Berkeley: University of California Press.
Holes, Clive. 1990. *Gulf Arabic*. Croom Helm Descriptive Grammars. London and New York: Routledge.
Holisky, Dee Ann and Rusudan Gagua. 1994. 'Tsova-Tush (Batsbi)'. In Rieks Smeets, ed., *The Indigenous Languages of the Caucasus, Volume 4: The North East Caucasian Languages*, pp. 147–212. Delmar, New York: Caravan Books.
Holst, Jan Henrik. 2001. *Lettische Grammatik*. Hamburg: Buske.
Hopper, Paul and Sandra A. Thompson. 1980. 'Transitivity in Grammar and Discourse'. *Language* 56: 251–99.
Horton, A. E. 1949. *A Grammar of Luvale*. Bantu Grammatical Archives edited by C. M. Doke. Johannesburg: Witwatersrand University Press.
Hoskison, James T. 1983. 'A Grammar and Dictionary of the Gude Language'. The Ohio State University, Ph.D. Dissertation.
Hosokawa, Komei. 1991. 'The Yawuru Language of West Kimberley: A Meaning-Based Description'. Australian National University Ph.D. Dissertation.
Hovdhaugen, Even and Ulrike Mosel, eds. 1999. *Negation in Oceanic Languages*. Typological Studies. LINCOM Studies in Austronesian Languages 2. München: Lincom Europa.
Hualde, José Ignacio and Jon Ortiz de Urbina, eds. 2003. *A Grammar of Basque*. Berlin and New York: Mouton de Gruyter.
Huang, Chu-Ren and Kathleen Ahrens. 1999. 'The Function and Category of *gei* in Mandarin Ditransitive Constructions'. *Journal of Chinese Linguistics* 27.2: 1–26.
Huffman, M. and R. Krakow, eds. 1993. *Nasals, Nasalization and the Velum*. San Diego: Academic Press.
Hurford, James R. 1975. *The Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.
—— 1987. *Language and Number: The Emergence of a Cognitive System*. Oxford: Blackwell.
Hurvich, L. M. 1982. 'Chromatic and Achromatic Response Functions'. In L. M. Hurvich, ed., *Color Vision*, pp. 52–76. Sunderland, MA: Sinauer.
Hutchison, John P. 1981. *The Kanuri Language: A Reference Grammar*. Madison: African Studies Program, University of Wisconsin.
Hutchisson, Don. 1986. 'Sursurunga Pronouns and the Special Uses of Quadral Number'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 1–20. Tübingen: Gunter Narr.
Huttar, George L. and Mary Huttar. 1994. *Ndyuka*. London: Routledge.
Hyman, Larry M. 1981. *Noni Grammatical Structure with Special Reference to Verb Morphology*. Southern California Occasional Papers in Linguistics 9. Los Angeles: Department of Linguistics, University of Southern California.
—— 2001. *Compositional and/or Templatic Morphology: Suffix Ordering in Bantu*. Paper presented at the Bay Area Typology Workshop, UC Berkeley, March 2001.

I

Ikoro, Suanu M. 1996. *The Kana Language*. Leiden: CNWS Research School.
Ingham, Bruce. 2001. *English-Lakota Dictionary*. Richmond, England: Curzon.

Innes, Gordon. 1966. *An Introduction to Grebo*. London: School of Oriental and African Studies, University of London.
Ittmann, Johannes. 1939. *Grammatik des Duala*. Berlin: Reimer.

J

Jacob, Judith M. 1968. *Introduction to Cambodian*. London: Oxford University Press.
Jacobsen, William. 1964. *A Grammar of the Washo Language*. University of California at Berkeley Dissertation.
Jacobson, Steven A. 1995. *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*. Fairbanks: Alaska Native Language Center.
Jacq, Pascale, and Paul Sidwell. 1999. *Sapuan*. München: Lincom Europa.
Jakobi, Angelika. 1990. *A Fur Grammar: Phonology, Morphophonology, and Morphology*. Hamburg: Helmut Buske.
Jakobson, Roman O. 1936 [1984]. 'Contribution to the General Theory of Case: General Meanings of the Russian Cases' [Translation of: Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutung der russischen Kasus]. In Linda R. Waugh and Morris Halle, eds., *Roman Jakobson, Russian and Slavic Grammar: Studies 1931–1981*, pp. 59–103. Berlin: Mouton de Gruyter.
—— 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.
Jakovlev N. F. 1940. *Sintakisis čečenskogo literaturnogo jazyka*. Moscow and Leningrad: Izdatel'stvo AN SSSR.
Jameson, Dorothea and L. M. Hurvich. 1968. 'Opponent-Response Functions Related to Measured Cone Photo-Pigments'. *Journal of the Optical Society of America* 58: 429–30.
Jenewari, Charles E. W. 1983. 'Defaka, Ijo's Closest Linguistic Relative'. In Ivan R. Dirhoff, ed., *Current Approaches to African Linguistics I*, Publications in African Languages and Linguistics 1, pp. 85–111. Dordrecht: Foris.
Jensen, C. 1998. *Comparative Tupi-Guarani Syntax*. Berlin: Mouton de Gruyter.
Jespersen, Otto. 1965. *The Philosophy of Language*. New York: Bantam.
Jochelson, W. 1905. 'Essay on the Grammar of the Yukaghir Language'. *American Anthropologist* 7.2: 369–424.
Johanson, Lars. 2000. 'Turkic Indirectives'. In Lars Johanson and Bo Utas, eds., *Evidentials: Turkic, Iranian and Neighbouring Languages*. Empirical Approaches to Language Typology 24, pp. 61–88. Berlin: Mouton de Gruyter.
Johanson, Lars and Bo Utas, eds. 2000. *Evidentials: Turkic, Iranian and Neighbouring Languages*. Empirical Approaches to Language Typology 24. Berlin: Mouton de Gruyter.
Johnson, Frederick and A. C. Madan. 1991 [1939]. *A Standard Swahili-English Dictionary*. Oxford: Oxford University Press.
Jones, Linda K. 1986. 'The Question of Ergativity in Yawa, a Papuan Language'. *Australian Journal of Linguistics* 6: 37–56.
Jones, Morris and Alan R. Thomas. 1977. *The Welsh Language: Studies in its Syntax and Semantics*. Cardiff: University of Wales Press.
Jones, Ross McCallum. 1998. *The Boko/Busa Language Cluster*. München: Lincom Europa.
Jones, Wendell and Paula Jones. 1991. *Barasano Syntax*. Publications in Linguistics 101. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive*. Cambridge: Cambridge University Press.
—— and Irene Philippaki-Warburton. 1987. *Modern Greek*. London and Wolfeboro, NH: Croom Helm.
—— —— 1989. *Modern Greek*. Croom Helm Descriptive Grammars Series. London and New York: Routledge.
Josephs, Lewis S. 1975. *Palauan Reference Grammar*. Honolulu: University Press of Hawaii.
Juldašev, A. A., ed. 1981. *Grammatika sovremennogo baškirskogo literaturnogo jazyka*. Moskva: Nauka.
Jung, Ingrid. 1989. 'Grammatik des Paez: ein Abriss'. University of Osnabrück Ph.D. Dissertation.

K

Kager, René. 1993. 'Alternatives to the Iambic-Trochaic Law'. *Natural Language and Linguistic Theory* 11: 381–432.
Kähler, Hans. 1963. *Texte von der Insel Simalur*. Berlin: D. Reimer.
Kahombo, Mateene. 1992. *Essai de grammaire du Kihunde*. Hamburger Beiträge zur Afrikanistik 1. Münster: Lit.
Kahrel, Peter. 1996. 'Aspects of Negation'. University of Amsterdam Ph.D. Dissertation.
Kahrel, Peter and René van den Berg, eds. 1994. *Typological Studies in Negation*. Typological Studies in Language 29. Amsterdam and Philadelphia: Benjamins.
Kakumasu, James. 1966. 'Urubú Sign Language'. *International Journal of American Linguistics* 34.4: 275–81.
—— 1986. 'Urubu-Kaapor'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 1*, pp. 326–403. Berlin: Mouton de Gruyter.
Kalinina, Elena. 1993. 'Sentences with Non-Verbal Predicates in the Sogratl Dialect of Avar'. In A. E. Kibrik, ed., *The Noun Phrase in the Andalal Dialect of Avar as Spoken at Sogratl*, pp. 90–104. EUROTYP Working Papers. Strasbourg: European Science Foundation.
Kalmár, Ivan. 1979. 'The Antipassive and Grammatical Relations in Eskimo'. In Frans Plank, ed., *Ergativity: Towards a Theory of Grammatical Relations*, pp. 117–43. London: Academic Press.
Kamba Muzenga, J. G. 1980. *Esquisse de grammaire Kete*. Annalen Menselijke Wetenschappen 104. Tervuren: Koninklijk Museum voor Midden-Afrika.
—— 1981. *Les formes verbales négatives dans les langues bantoues*. Annalen, Menselijke Wetenschappen 106. Tervuren: Koninklijk Museum voor Midden-Afrika.
Karlsson, Fred. 1978. *Finsk Grammatik*. Helsinki: SKS.
Kay, Paul and Brent Berlin. 1997. 'Science ≠ Imperialism: There are nontrivial constraints on color naming. Peer Commentary on Saunders and van Brakel'. *Behavioral and Brain Sciences* 20: 196–201.
—— —— Luisa Maffi and William Merrifield. 1997. 'Color Naming across Languages'. In C. L. Hardin and Luisa Maffi, eds., *Color Categories in Thought and Language*, pp. 21–56. Cambridge: Cambridge University Press.
—— —— and William Merrifield. 1991. 'Biocultural Implications of Systems of Color Naming'. *Journal of Linguistic Anthropology* 1: 12–25.
—— and Luisa Maffi. 1999. 'Color Appearance and the Emergence and Evolution of Basic Color Lexicons'. *American Anthropologist* 101: 743–60.
—— and Chad K. McDaniel. 1978. 'The Linguistic Significance of the Meanings of Basic Color Terms'. *Language* 54: 610–46.
Kazenin, Konstantin I. 1996. 'The Syntactic Structure of Action Nominal Constructions'. In A.E. Kibrik, ed., *Godoberi*, pp. 160–1. München: Lincom Europa.
—— (forthcoming). 'Reciprocal, Comitative, Sociative and Reflexive in Kabardian'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Keegan, John M. 1997. *A Reference Grammar of Mbay*. München: Lincom Europa.
Keen, Sandra. 1983. 'Yukulta'. In R. M. W. Dixon and Barry J. Blake, eds., *Handbook of Australian Languages 3*, pp. 191–304. Amsterdam: John Benjamins.
Keenan, Edward L. 1976. 'Remarkable Subjects in Malagasy'. In Charles Li, ed., *Subject and Topic*, pp. 247–301. New York: Academic Press.
—— and Bernand Comrie. 1977. 'Noun phrase accessibility and universal grammar'. *Linguistic Inquiry* 8: 63–99.
—— and Leonard M. Faltz. 1986. *Boolean Semantics for Natural Language*. Dordrecht: Reidel.
Kemmer, Suzanne. 1993. *The Middle Voice*. Typological Studies in Language 23. Amsterdam and Philadelphia: John Benjamins.
Kemmer, Suzanne. 1996. 'Collective and Distributive Marking, or: Where Unity meets Multiplicity'. In Alan K. Melby, ed., *The twenty-third LACUS Forum 1996*, pp. 231–49. Chapel Hill, NC: Linguistics Association of Canada and the United States.
Kenesei, István, Robert M. Vago, and Anna Fenyvesi. 1998. *Hungarian*. Descriptive Grammars. London and New York: Routledge.
Kenstowicz, M. 1994. 'Sonority Driven Stress'. Manuscript, MIT.
Key, Harold H. 1967. *Morphology of Cayuvava*. The Hague: Mouton.
Kibrik, Aleksandr E. 1972. 'O formal'nom vydelenii soglasovatel'nyx klassov v arčinskom jazyke'. *Voprosy jazykoznanija* 1: 124–31.
—— 1981. *Materialy k tipologii èrgativnosti: Beztinskij jazyk*. Moscow: Institut russkogo jazyka AN SSSR.
—— 1994. 'Archi'. In Rieks Smeets, ed., *The Indigenous Languages of the Caucasus 4*, Part 2, pp. 297–365. Delmar, New York: Caravan Books.
—— S. V. Kodzasov, I. P. Olovjannikova and D. S. Samedov. 1977. *Opyt strukturnogo opisanija archinskogo jazyka, volume 1–4*. Moscow: Izdatel'stvo Moskovskogo Universiteta.
Kibrik, Andrej. 1992. 'Relativization in polysynthetic languages'. *International Journal of American Linguistics* 58. 2: 135–157.
Kießling, Roland. 1994. *Eine Grammatik des Burunge*. Afrikanistische Forschungen 13. Hamburg: Research and Progress Verlag.
Kilian-Hatz, Christa. 1995. *Das Baka: Grundzüge einer Grammatik aus der Grammatikalisierungsperspektive*. Afrikanistische Monographien 6. Köln: Institut fur Afrikanistik, Universität zu Köln.

Kilpatrick, Eileen. 1981. 'Avokaya Predication'. *Occasional Papers in the Study of Sudanese Languages* 1: 91–109.

Kimball, Geoffrey D. 1988. 'Koasati Reduplication'. In William Shiplay, ed., *In Honor of Mary Haas: from the Haas Festival Conference on Native American Linguistics*, pp. 431–42. Berlin: Mouton de Gruyter.

—— 1991. *Koasati Grammar*. Studies in the Anthropology of North American Indians. Lincoln and London: University of Nebraska Press.

Kimenyi, Alexandre. 1980. *A Relational Grammar of Kinyarwanda*. Berkeley: University of California Press.

—— 1988. 'Passives in Kinyarwanda'. In Masayoshi Shibatani, ed., *Passive and Voice*, pp. 355–86. Amsterdam: John Benjamins.

King, Gareth. 1993. *Modern Welsh. A Comprehensive Grammar*. London and New York: Routledge.

Klein, Wolfgang. 1994. *Time in Language*. London: Routledge.

Klima, Edward and Ursula Bellugi. 1979. *The Signs of Language*. Cambridge, MA: Harvard University Press.

Klimov, Georgij Andreevič. 1970. *Jazyk burušaski*. Moscow: Nauka.

Koch, Harold. 1995. 'The Creation of Morphological Zeroes'. In G. Booij and J. van Marle, eds., *Yearbook of Morphology 1994*, pp. 31–71. Dordrecht: Kluwer.

—— 1997. *Morphological Changes in Arandic Pronouns*. Paper presented at the XIII International Conference on Historical Linguistics, Düsseldorf, August 15, 1997.

Koehn, Edward and Sally Koehn. 1986. 'Apalaí'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 1*, pp. 33–127. Berlin: Mouton de Gruyter.

Koelle, S. W. 1854. *Outlines of a Grammar of the Vei Language*. London: Kegan Paul.

Kohnen, B. 1933. *Shilluk Grammar, with a little English-Shilluk Dictionary*. Verona: Missioni Africane.

Kolia, J. A. 1975. 'A Balawaia Grammar Sketch and Vocabulary'. In T. E. Dutton, ed., *Studies in Languages of Central and Southeast Papua*, Pacific Linguistics, Series C-29, pp. 107–226. Canberra: Australian National University.

König, Ekkehard. 1991. *The Meaning of Focus Particles: A Comparative Perspective*. London and New York: Routledge.

—— and Peter Siemund. 2000a. 'Intensifiers and Reflexives: A Typological Perspective'. In Zygmunt Frajzyngier and Traci S. Curl, eds., *Reflexives: Forms and Functions*, pp. 41–74. Amsterdam: John Benjamins.

—— —— 2000b. 'The Development of Complex Reflexives and Intensifiers in English'. *Diachronica* 17: 39–84.

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.

—— 1994. 'Finiteness'. In R. E. Asher and J. M. Simpson, eds., *Encyclopedia of Language and Linguistics*, pp. 1245–8. Oxford and Aberdeen: Pergamon Press and Aberdeen University Press.

—— 2002. 'Action Nominal Constructions in the Languages of Europe'. In Frans Plank, ed., *The Noun Phrase in the Languages of Europe*, pp. 723–59. Berlin: Mouton de Gruyter.

Kordić, Sanježana. 1997. *Serbo-Croatian*. München: Lincom Europa.

Kornfilt, Jaklin. 1997. *Turkish*. Descriptive Grammars. London: Routledge.

Korolev, Nikolaj I. 1965. *Jazyk nepali*. Moscow: Nauka.

Kortmann, Bernd. 1997. *Adverbial Subordination: A Typology and History of Adverbial Subordinators Based on European Languages*. Berlin: Mouton de Gruyter.

Koshal, Sanyukta. 1979. *Ladakhi Grammar*. Delhi: Motilal Banarsidass.

Kossmann, Maarten G. 2000. *Esquisse grammaticale du rifain oriental*. Paris: Peeters.

Koval', A. I. 1979. 'O značenii morfologičeskogo pokazatelja klassa v fula'. In N. V. Oxotina, ed., *Morfonologija i morfologija klassov slov v jazykax Afriki*, pp. 5–100. Moscow: Nauka.

Kozinsky, Isaac, Vladimir Nedjalkov, and Maria Polinskaja. 1988. 'Antipassive in Chukchee'. In Masayoshi Shibatani, ed., *Passive and Voice*, pp. 651–706. Amsterdam: John Benjamins.

Kozintseva, Natalia A. 1995. *Modern Eastern Armenian*. München: Lincom Europa.

—— 2001. 'Imperative Constructions in Armenian'. In Victor S. Xrakovskij, ed., *Typology of Imperative Constructions*, pp. 245–67. München: Lincom Europa.

Kroeber, Alfred Louis. 1907. *The Washo Language of East Central California and Nevada*. University of California Publications in American Archaeology and Ethnology 4.5. Berkeley, CA: University of California Press.

—— and George William Grace. 1960. *The Sparkman Grammar of Luiseño*. University of California Publications in Linguistics 16. Berkeley and Los Angeles: University of California Press.

Kroeber, Paul Demarest. 1986. 'Antipassives and the Differentiation of Progressives Aspect in Southern Interior Salish'. *Chicago Linguistic Society* 22.1: 75–89.

Kroeger, Paul R. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI.

Kruspe, Nicole. 1999. *A Grammar of Semelai*. University of Melbourne Ph.D. Dissertation.

—— 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press.

Kühner, Raphael and Carl Stegmann. 1955. *Ausführliche Grammatik der lateinischen Sprache: Satzlehre*. Leverkusen: Gottschalk.

Kuipers, Aert H. 1967. *The Squamish Language, Grammar, Texts, Dictionary*. Janua Linguarum, Series Practica 73. The Hague: Mouton.

—— 1974. *The Shuswap Language: Grammar, Texts, Dictionary*. The Hague: Mouton.

Kuno, Susumo. 1973. *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

—— 1978. 'Japanese: A Characteristic OV Language'. In Winfred P. Lehmann, ed., *Syntactic Typology*, pp. 57–138. Austin: University of Texas Press.

Kunz, George Frederick. 1917. *Rings for the Finger*. Philadelphia: J. B. Lippincott Company.

Kuteva, Tania and Bernard Comrie. (forthcoming.) 'On relative clause formation in African languages'. MS. Max Planck Institute for Evolutionary Anthropology.

Kutsch Lojenga, Constance. 1994. *Ngiti: a Central Sudanic Language of Zaire*. Nilo-Saharan Linguistic Analyses and Documentation 9. Köln: Rüdiger Köppe.

Kuular Klara B. (forthcoming.) 'Reciprocals, Sociatives, Comitatives and Assistives in Tuva'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

Kuzmenkov Evgenij A., Natalija S. Yakhontova, Elena Skribnik, and Vladimir P. Nedjalkov. (forthcoming.) 'Reciprocal, Sociative and Assistive Constructions in Mongolian and Buryat'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

Kwee, John B. 1981. *Teach yourself Indonesian. A Complete Working Course*. London: Hodder and Stoughton.

L

Laidig, Wyn D. 1993. 'Insights from Larike Possessive Constructions'. *Oceanic Linguistics* 32: 312–51.

Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lambdin, Thomas O. 1983. *Introduction to Sahidic Coptic*. Macon, Georgia: Mercer University Press.

Lang, Adrianne. 1973. *Enga Dictionary with English Index*. Pacific Linguistics, Series C-20. Canberra: Australian National University.

Langacker, R. W. 1987a. *Foundations of Cognitive Grammar. Volume 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

—— 1987b. 'Nouns and Verbs'. *Language* 63: 53–64

—— 1991. *Foundations of Cognitive Grammar. Volume 2: Descriptive Applications*. Stanford: Stanford University Press.

Langdon, Margaret H. 1970. *A Grammar of Diegueño: The Mesa Grande Dialect*. University of California Publications in Linguistics 66. Berkeley and Los Angeles: University of California Press.

Last, Marco and Deborah Lucassen. 1998. 'A Grammatical Sketch of Chai, a Southeastern Surmic Language'. In Gerrit J. Dimmendaal and Marco Last, eds., *Surmic Languages and Cultures*, pp. 375–436. Köln: Rüdiger Köppe.

Lastra de Suárez, Yolanda. 1984. 'Chichimeco Jonaz'. In Munro Edmonson, ed., *Supplement to the Handbook of Middle American Languages: Linguistics*, pp. 20–42. Austin: University of Texas Press.

Lavotha, Ödön. 1973. *Kurzgefaßte estnische Grammatik*. Wiesbaden: Harrassowitz.

Lawrence, Marshall. 1972. 'Oksapmin Sentence Structure'. In *Papers in New Guinea Linguistics* 16, Pacific Linguistics, Series A-34, pp. 17–46. Canberra: Australian National University.

Lawrence, Marshall. 1993. *Oksapmin Dictionary*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Laycock, D. C. 1965. *The Ndu Language Family*. Linguistic Circle of Canberra Publications. Canberra: Australian National University.

Leavitt, Robert M. 1996. *Passamaquoddy-Maliseet*. Languages of the World/Materials 27. München: Lincom Europa.

Leeding, Velma J. 1989. 'Anindilyakwa Phonology and Morphology'. University of Sydney Ph.D. Dissertation.

Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Gunter Narr.

Leinonen, Marja. 2000. 'Evidentiality in Komi Zyryan'. In Lars Johanson and Bo Utas, eds., *Evidentials*, pp. 419–40. Berlin: Mouton de Gruyter.
Leslau, Wolf. 1995. *Reference Grammar of Amharic*. Wiesbaden: Harassowitz.
Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
Levinsohn, Stephen H. 1976. *The Inga Language*. The Hague: Mouton.
Lewis, Geoffrey L. 1967. *Turkish Grammar*. Oxford: Clarendon Press.
Lewis, Grout. 1893. *The Isizulu: a Revised Edition of a Grammar of the Zulu Language*. With an introduction and appendix. London: Kegan Paul.
Lewis, M. B. 1968. *Malay*. London: St. Paul's House.
Li, Charles N. and Sandra A. Thompson. 1973. 'Serial Verb Constructions in Mandarin Chinese: Coordination or Subordination? In You'll Take the High Node and I'll Take the Low Node'. Papers from the Comparative Syntax Festival, *Chicago Linguistics Society* 9: 96–113.
—— —— 1977. 'A Mechanism for the Development of Copula Morphemes'. In Charles N. Li, ed., *Mechanisms of Syntactic Change*, pp. 419–44. Austin: University of Texas Press.
—— —— 1981. *Mandarin Chinese: a Functional Reference Grammar*. Berkeley: University of California Press.
Lichtenberk, Frantisek. 1983a. *A Grammar of Manam*. Honolulu: University of Hawaii Press.
—— 1983b. Relational Classifiers. *Lingua 60*: 147–76.
—— 1985. 'Multiple Uses of Reciprocal Constructions'. *Australian Journal of Linguistics* 5: 19–41.
—— 1999. 'Reciprocals without Reflexives'. In Zygmunt Frajzyngier and Traci S. Curl, eds., *Reciprocals: Forms and Functions*, Typological Studies in Language 41, pp. 31–62. Amsterdam: John Benjamins.
—— (forthcoming.) 'Reciprocals and related meanings in To'aba'ita'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Lidz, Jeffrey. 1996. *Dimensions of Reflexivity*. University of Delaware Ph.D. Dissertation.
Lin, Xiangrong. 1993. *Jiarongyu yanjiu*. Chengdu: Sichuan Minzu Chubanshe.
Lindblom, Björn and Ian Maddieson. 1988. 'Phonetic Universals in Consonant Systems'. In C. Li and L. M. Hyman, eds., *Language, Speech and Mind*, pp. 62–78. London: Routledge.
Lindstedt, Jouko. 2000. 'The Perfect—Aspectual, Temporal and Evidential'. In Östen Dahl, ed., *Tense and Aspect in the Languages of Europe*, pp. 365–84. Berlin: Mouton de Gruyter.
Link, Godehard. 1986a. *Generalized Quantifiers and Plurals*. Report No. CSLI-86-87. Stanford: Center for the Study of Language and Information.
—— 1986b. *Je drei Äpfel—three apples each. Quantification and the German je*. MS. Stanford: Center for the Study of Language and Information.
—— 1991. 'Quantity and Number'. In D. Zaefferer, ed., *Semantic Universals and Universal Semantics*, pp. 133–49. Berlin: Foris Press.
Lister-Turner, R. and J. B. Clark. 1930. *Revised Motu Grammar and Vocabulary*. Edited by Percy Chatterton. Port Moresby: Government Printer.
Loeweke, Eunice and Jean May. 1980. 'General Grammar of Fasu (Namo Me)'. In Don Hutchisson, ed., *Grammatical Studies in Fasu and Mt. Koiali*, Workpapers in Papua New Guinea Languages 27, pp. 5–106. Ukarumpa: Summer Institute of Linguistics.
Lomtatidze, K. V. 1996. *Kategorija vzaimnosti (vzaimosojuznosti, vzaimoobojudnosti) v abxazo-adygskix jazykax*. Tbilisi: AN SSSR.
Loogman, Alfons. 1965. *Swahili Grammar and Syntax*. Pittsburgh: Duquesne University Press.
Lorimer, D. L. R. 1935. *The Burushaski Language. Volume 1: Introduction and Grammar*. Instituttet for Sammenlignende Kulturforskning, Serie B, Skrifter XXIX. Cambridge, MA: Harvard University Press/Oslo: Aschehoug.
Louw, J. A. 1975. 'Clicks as Loans in Xhosa'. In A. Traill, ed., *Bushman and Hottentot Linguistic Studies*, pp. 82–100. Johannesburg: African Studies Institute, University of the Witwatersrand.
Love, J. R. B. 2000. *The Grammatical Structure of the Worora Language of North-Western Australia*. München: Lincom Europa.
Lukas, Johannes. 1953. *Die Sprache der Tubu in der zentralen Sahara*. Berlin: Akademie-Verlag.
Lyman, Thomas Ammis. 1979. *Grammar of Mong Njua (Green Miao): A Descriptive Linguistic Study*. Sattley, CA: Blue Oak Press.
Lynch, John. 1982. 'Anejom Grammar Sketch'. In *Papers in Linguistics in Melanesia 4*, Pacific Linguistics, Series A-64, pp. 93–154. Canberra: Australian National University.
Lyons, John. 1968. *Introduction to Theoretical Linguistics*. London: Cambridge University Press.

M

Macalister, R. A. Stewart. 1914. *The Language of the Nawar or Zutt, the Nomad Smiths of Palestine*. Gypsy Lore Society Monograph 3. London: Bernard Quaritch.
Macaulay, Monica. 1992. 'Inverse Marking in Karuk: the Function of the Suffix *-ap*'. *International Journal of American Linguistics* 58.2: 182–201.
—— 1996. *A Grammar of Chalcatongo Mixtec*. University of California Publications in Linguistics 127. Berkeley: University of California Press.
MacDonald, Lorna. 1990. *A Grammar of Tauya*. Berlin: Mouton de Gruyter.
Machobane, Malillo. 1989. 'Some Restrictions on the Sesotho Transitivizing Morphemes'. McGill University Ph.D. Dissertation.
MacKay, Carolyn Joyce. 1999. *A Grammar of Misantla Totanac*. Salt Lake City: University of Utah Press.
Maddieson, Ian, C. L. Smith and N. Bessell. 2001. 'Aspects of the Phonetics of Tlingit'. *Anthropological Linguistics* 43: 135–176.
Mahootian, Shahrzad. 1997. *Persian*. London: Routledge.
Maiden, Martin and Cecilia Robustelli. 2000. *A Reference Grammar of Italian*. Chicago: NTC Publishing.
Majtinskaja K. E. 1969. *Mestoimenija v jazykax raznyx sistem*. Moscow: Nauka.
Makino, S. 1968. 'Japanese 'be'.' In John W.M. Verhaar (ed.) *The Verb 'Be' and its synonyms*. Part 2. Dordrecht: Reidel, 1–19.
Malchukov, Andrej L. (forthcoming.) 'Reciprocal and Sociative Constructions in Even'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Malsch, Derry L. 1987. 'The Grammaticalization of Social Relationship: the Origin of Number to Encode Deference'. In Anna Giacalone Ramat, Onofrio Carrubo, and Giuliano Bernini, eds., *Papers from the 7th International Conference on Historical Linguistics*, pp. 405–18. Amsterdam: John Benjamins.
Marchese, Lynell. 1986. 'The Pronominal System of Godie'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 217–55. Tübingen: Gunter Narr.
Maring, Joel M. 1967. 'Grammar of Acoma Keresan'. Indiana University Ph.D. Dissertation.
Markey, T. L. 1985. 'On Suppletion'. *Diachronica* 11: 51–66.
Maslova, Elena. 1999. 'Reciprocals and Set Construal'. In Zygmunt Frajzyngier and Traci S. Curl, eds., *Reciprocals: Forms and Functions*, Typological Studies in Language 41, pp. 161–79. Amsterdam: John Benjamins.
—— 2003. *A Grammar of Kolyma Yukaghir*. Berlin and New York: Mouton de Gruyter.
Mateene, Kahombo. 1992. *Essai de grammaire du Kihunde*. Hamburger Beiträge zur Afrikanistik 1. Münster and Hamburg: Lit.
Mathiassen, Terje. 1997. *A Short Grammar of Latvian*. Columbus, Ohio: Slavica Publishers.
Mathiot, Madeleine. 1962. 'Noun Classes and Folk Taxonomy in Papago'. *American Anthropologist* 64: 340–50. [Reprinted in Dell H. Hymes, ed. 1964. *Language in Culture and Society: A Reader in Linguistics and Anthropology*. New York: Harper & Row.]
Matisoff, James A. 1975. 'Rhinoglottophilia: the Mysterious Connection between Nasality and Glottality'. In Charles A. Ferguson, Larry M. Hyman, and John Ohala, eds., *Nasàlfest*, pp. 265–88. Stanford: Language Universals Project, Department of Linguistics, Stanford University.
Matthews, Peter H. 1974. *Morphology: An Introduction to the Theory of Word Structure*. Cambridge: Cambridge University Press.
Matthews, Stephen and Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. New York: Routledge.
Matthewson, Lisa. 1998. *Determiner Systems and Quantificational Strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.
McClelland, Clive W. 2000. *The Interrelations of Syntax, Narrative Structure, and Prosody in a Berber Language*. Lewiston: Edwin Mellen Press.
McGregor, R. S. 1977. *Outline of Hindi Grammar*. Delhi: Oxford University Press. [2nd edn.]
—— 1995. *Outline of Hindi Grammar: with Exercises*. Oxford: Oxford University Press. [3rd edn.]
McGregor, William B. 1990. *A Functional Grammar of Gooniyandi*. Studies in Language Companion Series 22. Amsterdam and Philadelphia: John Benjamins.
McKay, Graham R. 1978. 'Pronominal Person and Number Categories in Rembarrnga and Djeebbana'. *Oceanic Linguistics* 17: 27–37.
McLendon, S. 1975. *A Grammar of Eastern Pomo*. Berkeley: University of California Press.
Meeuwis, Michael. 1998. *Lingala*. Languages of the World/Materials 261. München: Lincom Europa.

Meier, Paul, Inge Meier, and John Bendor-Samuel. 1975. *A Grammar of Izi: An Igbo Language*. Norman, Oklahoma: Summer Institute of Linguistics.
Meinhof, Carl. 1930. *Der Koranadialekt des Hottentottischen*. Beiheft zur Zeitschrift für Eingeborenensprachen 12. Berlin: Verlag Dietrich Reimer (Ernst Vohsen) A. G. [Reprinted 1969, Nendeln, Liechtenstein: Kraus.]
Mel'čuk, Igor. 1994. 'Suppletion: Toward a Logical Analysis of the Concept'. *Studies in Language* 18.2: 339–410.
Merlan, Francesca C. 1982. *Mangarayi*. Lingua Descriptive Studies 4. Amsterdam: North Holland Publishing Company.
—— 1983. *Ngalakan Grammar, Texts and Vocabulary*. Pacific Linguistics, Series B-89. Canberra: Australian National University.
—— 1985. 'Split Intransitivity'. In Johanna Nichols and Anthony C. Woodbury, eds., *Grammar Inside and Outside the Clause*, pp. 324–62. Cambridge: Cambridge University Press.
—— 1994. *A Grammar of Wardaman, a Language of the Northern Territory of Australia*. Mouton Grammar Library 11. Berlin and New York: Mouton de Gruyter.
Meyer-Bahlburg, Hilke. 1972. *Studien zur Morphologie und Syntax des Musgu*. Hamburg: Helmut Buske Verlag.
Miestamo, Matti. 2000. 'Towards a Typology of Standard Negation'. *Nordic Journal of Linguistics* 23: 65–88.
—— 2003. 'Clausal Negation: A Typological Study'. University of Helsinki Ph.D. Dissertation.
Miller, Amy W. 2001. *A Grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter.
Miller, Wick R. 1965. *Acoma Grammar and Texts*. University of California Publications in Linguistics 40. Berkeley: University of California Press.
Mills, Anne E. 1986. *The Acquisition of Gender: A Study of English and German*. Berlin: Springer Verlag.
Minassian, Martiros. 1980. *Grammaire d'arménien oriental*. Delmar, NY: Caravan Books.
Mintz, Malcolm W. 1994. *A Student's Grammar of Malay and Indonesian*. Singapore: EPB.
Miranda, Rocky V. 1975. 'Indo-European Gender: a Study in Semantic and Syntactic Change'. *Journal of Indo-European Studies* 3: 199–215.
Mirikitani, Leatrice T. 1972. *Kapampangan Syntax*. Oceanic Linguistics Special Publication 10. Honolulu: University Press of Hawaii.
Mitchell, T. F. 1962. *Colloquial Arabic: the Living Language of Egypt*. London: The English Universities Press.
Mithun, Marianne. 1988a. 'The Grammaticalization of Coordination'. In John Haiman and Sandra A. Thompson, eds., *Clause Combining in Grammar and Discourse*, pp. 331–59. Amsterdam: John Benjamins.
—— 1988b. 'Lexical Categories and the Evolution of Verbal Number'. In Hammond Michael and Michael Noonan, eds., *Theoretical Morphology*, pp. 211–33. San Diego: Academic Press.
—— 1991. 'Active/Agentive Case Marking and its Motivations'. *Language* 67: 510–46.
Moore, Denny. 1984. *Syntax of the Language of the Gavião Indians of Rondônia, Brazil*. City University of New York Ph.D. Dissertation.
Moravcsik, Edith A. 1978. 'Reduplicative Constructions'. In J. H. Greenberg et al., eds., *Universals of Human Language, Volume 3: Word Structure*, pp. 297–334. Stanford: Stanford University Press.
—— 2003. 'A Semantic Analysis of Associative Plurals'. *Studies in Language* 27.3: 469–503.
Morgenstierne, Georg. 1949. 'The Language of the Prasun Kafirs'. *Norsk Tidsskrift for Sprogvidenskap* 15: 186–334.
Mosel, Ulrike. 1984. *Tolai Syntax and its Historical Development*. Canberra: Australian National University.
—— and Ruth Spriggs. 2000. 'Gender in Teop (Bougainville, Papua New Guinea)'. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, and Mirja Saari, eds., *Gender in Grammar and Cognition*, Trends in Linguistics: Studies and Monographs 124, pp. 321–49. Berlin: Mouton de Gruyter.
Moseley, Christopher and R. E. Asher (eds.) 1994. *Atlas of the World's Languages*. London: Routledge.
Moshinsky, Julius. 1974. *A Grammar of Southeastern Pomo*. University of California Publications in Linguistics 72. Berkeley and Los Angeles: University of California Press.
Mous, Maarten. 1992. *A Grammar of Iraqw*. Rijksuniversiteit Leiden Doctoral Dissertation.
—— 1993. *A Grammar of Iraqw*. Kuschitische Sprachstudien 9. Hamburg: H. Buske.
Moyse-Faurie, Claire. 1983. *Le drehu, langue de Lifou (Îles Loyauté)*. Paris: SELAF.
—— (forthcoming.) 'Reciprocal, Sociative, Reflexive and Iterative Constructions in East Futunan'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Mühlhäusler, Peter and Rom Harré. 1990. *Pronouns and People. The Linguistic Construction of Social and Personal Identity*. Oxford: Blackwell.
Mulder, Jean Gail. 1994. *Ergativity in Coast Tsimshian (Sm'algyax)*. Berkeley: University of California Press.
Müller, Natascha. 2000. 'Gender and Number in Acquisition'. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, and Mirja Saari, eds., *Gender in Grammar and Cognition*, Trends in Linguistics: Studies and Monographs 124, pp. 351–99. Berlin: Mouton de Gruyter.
Munro, Pamela. 1982. 'On the Transitivity of "say" Verbs'. In P. Hopper and S. A. Thompson, eds., *Studies in Transitivity*, pp. 301–18. New York: Academic Press.
Munro, Pamela and Lynn Gordon. 1982. 'Syntactic Relations in Western Muskogean'. *Language* 58.1: 81–115.
Muntzel, Martha C. 1986. *The Structure of Ocuilteco*. State University of New York at Albany Ph.D. Dissertation.
Murane, Elizabeth. 1974. *Daga Grammar: From Morpheme to Discourse*. Summer Institute of Linguistics Publications in Linguistics and related Fields 43. Norman: Summer Institute of Linguistics.
Mushin, Ilana. 1995. Epistememes in Australian Languages. *Australian Journal of Linguistics* 15: 1–31.

N

Nagaraja, K. S. 1985. *Khasi: A Descriptive Study*. Pune: Deccan College.
—— 1999. *Korku Language: Grammar, Texts and Vocabulary*. Tokyo: Tokyo University of Foreign Studies.
Najlis, E. 1973. *Lengua selknam*. Filología y Lingüística 3. Buenos Aires: Universidad de Salvador.
Nebel, Arturo. 1948. *Dinka Grammar (Rek-Malual Dialect) with Texts and Vocabulary*. Verona: Missioni Africane.
Nedjalkov, Igor V. 1997. *Evenki*. Routledge Descriptive Grammars. London and New York: Routledge.
—— and Vladimir P. Nedjalkov. (forthcoming a.) 'Reciprocal and Sociative Constructions in Evenki'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— (forthcoming b.) 'Reciprocals, Sociatives, Comitatives and Assistives in Yakut'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Nedjalkov, Vladimir P. (forthcoming a.) 'Reciprocal Constructions in Chukchi (with an appendix on Koryak)'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— (forthcoming b.) 'Reciprocals, Assistives and Plural in Kirghiz'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Nedjalkov, Vladimir P., ed., with the assistance of Emma Geniušienė and Zlatka Guentchéva. (forthcoming.) *Typology of Reciprocal Constructions*. München: Lincom Europa.
Nedjalkov, Vladimir P. and Igor V. Nedjalkov. (forthcoming.) 'Reciprocals, Sociatives and Competitives in Karachai-Balkar'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— and G. G. Silnitsky. 1973. 'The Typology of Morphological and Lexical Causatives'. In Ferenc Kiefer, ed., *Trends in Soviet Theoretical Linguistics*, pp. 1–32. Dordrecht: Reidel.
Nettle, Daniel. 1998. *The Fyem Language of Northern Nigeria*. München: Lincom Europa.
—— 1999. *Linguistic Diversity*. Oxford: Oxford University Press.
Newell, Leonard E. 1993. *Batad Ifugao Dictionary: with Ethnographic Notes*. Manila: Linguistic Society of the Philippines.
Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven, CN: Yale University Press.
Newman, Stanley. 1944. *The Yokuts Language of California*. New York: Johnson Reprint Co.
Nguyên, Đinh-Hòa. 1966. *Vietnamese-English Dictionary*. Rutland, VT and Tokyo: Charles E. Tuttle & Co.
Nichols, Johanna. 1986. 'Head-Marking and Dependent-Marking Grammar'. *Language* 66: 56–119.
—— 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
—— 1994. 'Ingush'. In Rieks Smeets, ed., *The Indigenous Languages of the Caucasus 4*, pp. 79–145. Delmar, New York: Caravan Books.
—— 2001. 'Why "me" and "thee"?' In Laurel Brinton, ed., *Historical Linguistics 1999*, pp. 253–76. Amsterdam and Philadelphia: John Benjamins.
—— and David A. Peterson. 1996. 'The Amerind Personal Pronouns'. *Language* 72.2: 336–71.

Nichols, John. 1988. *An Ojibwe Text Anthology*. London and Ontario: Center for Research and Teaching of Canadian Native Languages.
Nida, Eugene. 1963. 'The Identification of Morphemes'. In Martin Joos, ed., *Readings in Linguistics 1*, pp. 255–71. New York: American Council of Learned Societies.
Niggemeyer, H. 1951–52. 'Alune-Sprache: Texte, Wörterverzeichnis und Grammatik einer Sprache West-Cerams'. *Zeitschrift für Ethnologie* 76: 50–69; 77: 116–32.
Nikolaeva, Irina. 1999. *Ostyak*. Languages of the World/Materials 305. München: Lincom Europa.
—— and Maria Tolskaya. 2001. *A Grammar of Udihe*. Berlin: Mouton de Gruyter.
—— and E. A. Xelimskij. 1997. 'Jukagirskij jazyk'. In A. P. Volodin et al., eds., *Jazyki mira: paleoaziatskie jazyki*, pp. 155–68. Moscow: Indrik.
Noonan, Michael. 1985. 'Complementation'. In T. Shopen, ed., *Language Typology and Syntactic Description, Volume 2: Complex Constructions*, pp. 42–140. Cambridge: Cambridge University Press.
—— 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.
—— 2000. *Recent Language Contact in the Nepal Himalaya*. Paper presented at the Sixth Himalayan Language Symposium, University of Wisconsin—Milwaukee, 15 June 2000. To appear in a volume edited by Randy LaPolla, Graham Thurgood, and Tasaku Tsunoda. Rough Draft downloadable under http://www.uwm.edu/ ~noonan/Papers.html
Nordlinger, Rachel. 1998. *A Grammar of Wambaya, Northern Territory (Australia)*. Pacific Linguistics, Series C-140. Canberra: Australian National University.
Norman, H. W., ed. 1848. *The Anglo-Saxon Version of the 'Hexameron' and the Anglo-Saxon Remains of St Basil's 'Admonitio ad Filium Spiritualem'*. London: Smith. [revised edn. 1849.]
Norris, Evan J. 1986. *A Grammar Sketch and Comparative Study of Eastern Mono*. University of California at San Diego Ph.D. Dissertation.
Noss, Richard B. 1964. *Thai Reference Grammar*. Washington: Foreign Service Institute, Dept. of State, United States Government.
Nougayrol, Pierre. 1990. *La langue des Aiki dits Rounga (Tchad, République centrafricaine): Esquisse descriptive et lexique*. Paris: Librairie Orientaliste Paul Geuthner.
Novikova, K. A. 1997. 'Èvenskij jazyk'. In V. M. Alpatov et al., eds., *Jazyki Mira: Mongol'skie jazyki, tunguso-man'čžurskie jazyki, japonskij jazyk, korejskij jazyk*, pp. 284–304. Moscow: Indrik.

O

Ó Siadhail, Michael. 1989. *Modern Irish: Grammatical Structure and Dialectal Variation.* Cambridge Studies in Linguistics—Supplementary Volume. Cambridge: Cambridge University Press.
Ogawa, A. 1998. 'Emphatische Reflexiva im Sprachvergleich: Deutsch-Japanisch'. *Sprachtypologie und Universalienforschung* 51: 164–78.
Ogloblin, Aleksandr A. and Vladimir P. Nedjalkov. (forthcoming.) 'Reciprocal constructions in Indonesian'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
O'Grady, Geoffrey N., Charles F. Voegelin, and F. M. Voegelin. 1966. 'Languages of the World: Indo-Pacific Fascicle six'. *Anthropological Linguistics* 8.2: 1–161.
Ogunbowale, P. O. 1970. *The Essentials of the Yoruba Language*. London: University of London Press.
Ohala, John J. 1983. 'The Origin of Sound Patterns in Vocal Tract Constraints'. In Peter F. MacNeilage, ed., *The Production of Speech*, pp. 189–216. Heidelberg and Berlin: Springer Verlag.
—— and M. Ohala. 1993. 'The Phonetics of Nasal Phonology: Theorems and Data'. In M. Huffman and R. Krakow, eds., *Nasals, Nasalization and the Velum*, pp. 225–50. San Diego: Academic Press.
Ohala, Manjari. 1983. *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.
Oinas, Felix J. 1966. *Basic Course in Estonian*. Bloomington: Indiana University.
Okell, John. 1969. *A Reference Grammar of Colloquial Burmese, Part 1*. London: Oxford University Press.
Olawsky, Knut J. 1999. *Aspects of Dagbani Grammar*. München: Lincom Europa.
Olson, Cliff. 1992. 'Gumawana (Amphlett Islands, Papua New Guinea): Grammar Sketch and Texts'. *Papers in Austronesian Linguistics* 2: 251–430.
Olzu Zubiri, Jesús and Miguel Angel Jusayú. 1985. *Gramática de la lengua guajira*. San Cristóbal: Universidad Católica de Tachira.
Onishi, Masayuki. 2000. 'Transitivity and Valency-Changing Derivations in Motuna'. In R. M. W. Dixon and Alexandra Y. Aikhenvald, eds., *Changing Valency*, pp. 115–44. Cambridge: Cambridge University Press.
Osada, Toshiki. 1992. *A Reference Grammar of Mundari*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
—— (forthcoming.) 'Reciprocals in Mundari'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Osborn, Henry A. Jr. 1966. 'Warao I: Phonology and Morphophonemics'. *International Journal of American Linguistics* 32: 108–23.
Osborne, C. R. 1974. *The Tiwi Language*. Australian Aboriginal Studies 55. Pacific Linguistics, Series 21. Canberra: Australian Institute of Aboriginal Studies.
Osumi, Midori. 1995. *Tinrin Grammar*. Honolulu: University of Hawaii Press.
Oswalt, Robert. 1991. 'A Method for Assessing Distant Linguistic Relationships'. In Sydney M. Lamb and E. Douglas Mitchell, eds., *Sprung from Some Common Source: Investigations into the Prehistory of Languages*, pp. 389–404. Stanford: Stanford University Press.
Otaina, Galina A. and Vladimir P. Nedjalkov. (forthcoming.) 'Reciprocal constructions in Nivkh (Gilyak)'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Kuschitische Sprachstudien, Cushitic Language Studies 4. Hamburg: Helmut Buske.

P

Palmer, Frank R. 1979. *Modality and the English Modals*. London: Longman.
Palmer, Frank R. 2001 [first pub. 1986]. *Mood and Modality*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press. [2nd edn.]
Pandharipande, Rajeshwari V. 1997. *Marathi*. London: Routledge.
Panfilov, V. Z. 1962. *Grammatika nivxskogo jazyka*. Moscow: Nauka.
Paris, Catherine. 1989. 'West Circassian (Adyghe: Abzakh dialect)'. In B. George Hewitt, ed., *The Indigenous Languages of the Caucasus, Volume 2: The Northwest Caucasian Languages*, pp. 155–260. Delmar, NY: Caravan Books.
Parker, Elizabeth. 1986. 'Mundani Pronouns'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 131–66. Tübingen: Gunter Narr.
Parker, E. M. and R. J. Hayward. 1985. *An Afar-English-French Dictionary (with Grammatical Notes in English)*. London: SOAS, University of London.
Paroz, R. A. 1946. *Elements of Southern Sotho*. Morija, Basutoland: Morija Sesuto Book Depot.
Payne, Doris L. 1990. *The Pragmatics of Word Order: Typological Dimensions of Verb Initial Languages*. Berlin: Mouton de Gruyter.
—— and David L. Payne. 1990. 'Yagua'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 2*, pp. 249–474. Berlin: Mouton de Gruyter.
Payne, J. R. 1979. 'The Decay of Ergativity in Pamir Languages'. *Lingua* 51: 147–86.
Pellizzaro, Siro M. 1969. *Shuar: Apuntes de gramática*. Quito: Offsetec.
Penchev, Jordan. (forthcoming.) 'Reciprocal and Reflexive Constructions in Bulgarian'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Penchoen, Thomas G. 1973. *Tamazight of the Ayt Ndhir.* Afroasiatic Dialects 1. Los Angeles: Undena Publishing.
Perkins, Michael R. 1983. *Modal Expressions in English*. Open Linguistics Series. London: Pinter.
Pet, Willem A. P. 1987. 'Lokono Dian, the Arawak Language of Suriname: A Sketch of its Grammatical Structure and Lexicon'. Cornell University Ph.D. Dissertation.
Peterson, David A. 1999. *Discourse-functional, Historical, and Typological Aspects of Applicative Constructions*. UC Berkeley Ph.D. Dissertation.
Petterson, Robert. 1999. *Rumu–English–Hiri-Motu Dictionary*. International Pacific College, Occasional Paper 6. Palmerston North, New Zealand: International Pacific College.
Pilhofer, Georg. 1933. *Grammatik der Kâte-Sprache in Neuguinea*. Zeitschrift für Eingeborenen-Sprachen, Beiheft 14. Berlin and Hamburg: D. Reimer (Ernst Vohsen). [Reprinted 1969 by Kraus Reprint, Nendeln, Liechtenstein.]
Pitkin, Harvey. 1985. *Wintu Grammar*. University of California Publications in Linguistics 94. Berkeley, CA: University of California Press.
Pitman, Donald. 1980. *Bosquejo de la gramática araona*. Notas lingüísticas 9. Riberalta, Bolivia: Instituto Lingüístico de Verano.
Plank, Frans, ed. 1995. *Double Case: Agreement by Suffixaufnahme*. Oxford and New York: Oxford University Press.
Plank, Frans. 1996. 'Domains of the Dual, in Maltese and in General'. *Rivista di Linguistica* 8.1: 123–40.

—— 1999. 'Split Morphology: how Agglutination and Flexion Mix'. *Linguistic Typology* 3: 279–340.

—— and Wolfgang Schellinger. 1997. 'The Uneven Distribution of Genders over Numbers: Greenberg Nos. 37 and 45'. *Linguistic Typology* 1.1: 53–101.

Plungian, Vladimir A. 2000. *Obščaja morfologija: Vvedenie v problematiku*. Moscow: URSS.

Polinsky, Maria. 1994. 'Relativization in Chukchee'. In Howard Aronson, ed., *Linguistic Studies in the Non-Slavic Languages of the Commonwealth of Independent States and the Baltic Republics*, pp. 241–62. Chicago: Chicago Linguistic Society.

—— 1995. 'Double Objects in Causatives'. *Studies in Language* 19: 121–219.

—— and Dan Jackson. 1999. 'Noun Classes: Language Change and Learning'. In Barbara A. Fox, Dan Jurafsky, and Laura A. Michaelis, eds., *Cognition and Function in Language*, pp. 29–50. Stanford: Center for the Study of Language and Information.

—— and Isaac Kozinsky. 1992. 'Ditransitive Constructions in Kinyarwanda'. *Papers from the 28th Regional Meeting of the Chicago Linguistic Society*, pp. 426–42. Chicago: Chicago Linguistic Society.

—— and Vladimir Nedjalkov. 1987. 'Contrasting the Absolutive in Chukchee: Syntax, Semantics, Pragmatics'. *Lingua* 71: 239–70.

—— and Ezra van Everbroeck. 2003. 'Development of Gender Classifications: Modeling the Historical Change from Latin to French'. *Language* 79: 356–90.

Popjes, Jack and Jo Popjes. 1986. 'Canela-Krahô'. In Desmond C. Derbyshire and Geoffrey K. Pullum, eds., *Handbook of Amazonian Languages 1*, pp. 128–99. Berlin: Mouton de Gruyter.

Popovich, Harold. 1986. 'The Nominal Reference System of Maxakalí'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 351–8. Tübingen: Gunter Narr.

Poppe, Nicholas. 1954. *Grammar of Written Mongolian*. Wiesbaden: Harrassowitz.

Portman, M. V. 1992 [1887]. *Andamanese Languages*. Delhi: Manas Publications.

Postal, Paul. 1977. 'Antipassive in French'. *Lingvisticae Investigationes* 1: 333–74.

Poulos, George. 1990. *A Linguistic Analysis of Venda*. Pretoria: Via Afrika Limited.

—— and Sonja E. Bosch. 1997. *Zulu*. Languages of the World/Materials 50. München: Lincom Europa.

Prasad, Bal Ram. 1991. *Mising Grammar*. Mysore, India: Central Institute of Indian Languages.

Press, Ian. 1986. *A Grammar of Modern Breton*. Berlin: Mouton de Gruyter.

Press, Margaret L. 1979. *Chemehuevi*. Berkeley: University of California Press.

Priego Montfort, Maria E. 1989. 'Gramática del Otomí (Hnahnu) del Mezquital, México'. University of Bielefeld Ph.D. Dissertation.

Primus, Beatrice. 1998. 'The Relative Order of Recipient and Patient in the Languages of Europe'. In Anna Siewierska, ed., *Constituent Order in the Languages of Europe*, pp. 421–73. Berlin: Mouton de Gruyter.

Prokof'eva, E. D. 1966. 'Sel'kupskij jazyk'. In *Jazyki narodov SSSR, Volume 3: Finno-ugorskie i samodijskie jazyki*, pp. 396–415. Moscow: Nauka.

Q

Quine, Willard van Orman. 1969. 'Ontological Relativity'. In Willard van Orman Quine, ed., *Ontological Relativity and Other Essays*, pp. 26–68. New York: Columbia University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

R

Rajaonarimanana, Narivelo. 1995. *Grammaire moderne de la langue malgache*. Paris: L'Asiathèque.

Randal, Scott. 1998. 'A Grammatical Sketch of Tennet'. In Gerrit J. Dimmendaal and Marco Last, eds., *Surmic Languages and Cultures*, pp. 219–72. Köln: Rüdiger Köppe.

Rau, Der-Hwa Victoria. 1992. 'A Grammar of Atayal'. Cornell University Ph.D. Dissertation. [Also published 1992 by Crane Publishing Co., Taipei.]

Raz, S. 1983. *Tigre Grammar and Texts*. Malibu: Undena.

Reed, Irene, Osahito Miyaoka, Steven Jacobson, Paschal Afcan, and Michael Krauss. 1977. *Yup'ik Eskimo Grammar*. Fairbanks: Alaska Native Language Center, University of Alaska.

Reed, Judy and David L. Payne. 1986. 'Asheninca (Campa) Pronominals'. In Wiesemann, Ursula, ed., *Pronominal Systems*, pp. 323–31. Tübingen: Gunter Narr.

Reesink, Ger P. 1999. *A Grammar of Hatam, Bird's Head Peninsula, Irian Jaya*. Pacific Linguistics, Series C-146. Canberra: Australian National University.

Refsing, Kirsten. 1986. *The Ainu Language: The Morphology and Syntax of the Shizunai Dialect*. Aarhus: Aarhus University Press.

Reh, Mechthild. 1985. *Die Krongo-Sprache (Nìinò Mó-Dì)*. Berlin: Dietrich Reimer.

—— 1996. *Anywa Language: Description and Internal Reconstructions*. Köln: Rüdiger Köppe.

Rehg, Kenneth L. 1981. *Ponapean Reference Grammar*. Honolulu: University Press of Hawaii.

Reichle, Verena. 1981. *Bawm Language and Lore*. Bern: Peter Lang.

Reid, Nicholas. 1997. 'Class and Classifier in Ngan'gityemerri'. In Mark Harvey and Nicholas Reid, eds., *Nominal Classification in Aboriginal Australia*, Studies in Language Companion Series 37, pp. 165–228. Amsterdam: John Benjamins.

Reinisch, Leo. 1893. *Die Bedauye-Sprache in Nord-Ost Afrika*. Wien: Tempsky.

Rennison, John R. 1997. *Koromfe*. London: Routledge.

Revithiadou, A. 1999. *Headmost Accent Wins: Head Dominance and Ideal Prosodic Form in Lexical Accent Systems*. HIL/LOT Dissertation 15. The Hague: Holland Academic Graphics.

Rice, Curt. 1992. 'Binarity and Ternarity in Metrical Theory: Parametric Extensions'. University of Texas, Austin Ph.D. Dissertation.

Rice, Keren. 1989. *A Grammar of Slave*. Mouton Grammar Library 5. Berlin: Mouton de Gruyter.

Ringe, Don. 1998. 'A Probabilistic Evaluation of Indo-Uralic'. In Joseph C. Salmons and Brian D. Joseph, eds., *Nostratic: Sifting the Evidence*, pp. 154–98. Amsterdam and Philadelphia: John Benjamins.

Rischel, Jørgen. 1974. *Topics in West Greenlandic Phonology*. Copenhagen: Akademisk Forlag.

Robbins, Frank E. 1968. *Quiotepec Chinantec Grammar*. Papeles de la Chinantla IV. México: Museo Nacional de Antropología.

Roberts, John R. 1987. *Amele*. Croom Helm Descriptive Grammar Series. London: Croom Helm.

Robins, R. H. 1958. *The Yurok Language, Grammar, Texts, Lexicon*. University of California Publications in Linguistics 15. Berkeley and Los Angeles: University of California Press.

Rodrigues, Aryon D. 1999. 'Macro-Jê'. In R. M. W. Dixon and Alexandra Y. Aikhenvald, eds., *The Amazonian Languages*, pp. 165–206. Cambridge: Cambridge University Press.

Rood, David S. 1976. *Wichita Grammar*. New York: Garland.

—— and Allan R. Taylor. 1996. 'Sketch of Lakhota, a Siouan Language'. In Ives Goddard, ed., *Handbook of North American Indians, Volume 17: Languages*, pp. 440–82. Washington: Smithsonian Institution.

Rosenbach, Anette. 2002. *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. Topics in English Linguistics 42. Berlin and New York: Walter de Gruyter.

Rosenthal, S. and H. van der Hulst. 1999. 'Weight by Position'. *Natural Language and Linguistic Theory* 17.3: 499–550.

Ross, John Robert. 1969. 'Auxiliaries as Main Verbs'. In William Todd, ed., *Studies in Philosophical Linguistics*, pp. 77–102. Evanstown, IL: Great Expectations.

Ross, Malcolm and John Natu Paol. 1978. *A Waskia Grammar Sketch and Vocabulary*. Canberra: Australian National University.

Rowlands, E. C. 1969. *Teach Yourself Yoruba*. London: English Universities Press.

Rubino, Carl. 1998. *Ilocano: Ilocano-English/English-Ilocano Dictionary and Phrasebook*. New York: Hippocrene Books.

—— 2000. *Ilocano Dictionary and Grammar*. Honolulu: University of Hawaii Press.

—— 2001. 'Pangasinan'. In Jane Garry and Carl Rubino, eds., *Encyclopedia of the World's Languages: Past and Present*, pp. 539–42. New York and Dublin: H. W. Wilson Press.

Rude, Noel E. 1985. 'Studies in Nez Perce Grammar and Discourse'. University of Oregon Ph.D. Dissertation.

Rudes, Blair. 1980. 'On the Nature of Verbal Suppletion'. *Linguistics* 18: 655–76.

Ruhlen, Merritt. 1975. 'The Patterning of Nasal Vowels'. In Charles A. Ferguson, Larry M. Hyman, and John Ohala, eds., *Nasàlfest*, pp. 333–52. Stanford: Language Universals Project, Department of Linguistics, Stanford University.

—— 1987. *A Guide to the World's Languages 1: Classification*. Stanford: Stanford University Press.

Rumsey, Alan. 1982. *An Intra-Sentence Grammar of Ungarinjin, North-Western Australia*. Pacific Linguistics, Series B-86. Canberra: Australian National University.
Rupp, James E. 1989. *Lealao Chinantec Syntax*. Studies in Chinantec Languages 2. Dallas, Texas: Summer Institute of Linguistics and the University of Texas at Arlington.
Russell, Kevin. 1999. 'The "Word" in two Polysynthetic Languages'. In T. Alan Hall and Ursula Kleinhenz, eds., *Studies on the Phonological Word*, pp. 203–21. Amsterdam: John Benjamins.

S

Sadock, Jerrold M. and Arnold M. Zwicky. 1985. 'Speech Acts Distinctions in Syntax'. In Timothy Shopen, ed., *Language Typology and Syntactic Description*, pp. 155–96. Cambridge: Cambridge University Press.
Saeed, John I. 1993. *Somali Reference Grammar*. Kensington, Maryland: Dunwoody Press.
Salas, Adalberto. 1992. *El mapuche o araucano: fonología, gramática, y antología de cuentos*. Colección Lenguas y Literaturas Indígenas 3. Madrid: Editorial MAPFRE.
Saltarelli, Mario, with Miren Azkarate, David Farwell, Jon Ortiz de Urbina, and Lourdes Oñederra. 1988. *Basque*. Croom Helm Descriptive Grammars. London: Croom Helm.
Samarin, William J. 1967. *A Grammar of Sango*. The Hague: Mouton.
Sammallahti, Pekka. 1998. 'Saamic'. In Daniel Abondolo, ed., *The Uralic Languages*, pp. 43–95. London and New York: Routledge.
Sanchez, M. and O. Castro. 1977. *Una gramática pedagógica del waunana*. Lenguas de Panama 3. Panama: Instituto Nacional de Cultura and Instituto Lingüístico de Verano.
Sapir, Edward. 1922. 'The Takelma Language of Southwestern Oregon'. In Franz Boas, ed., *Handbook of American Indian Languages 2*, Smithsonian Institution Bureau of American Ethnology Bulletin 40, pp. 1–296. Washington: Government Printing Office.
Sapir, J. David. 1965. *A Grammar of Diola-Fogny*. West African Language Monographs 3. Ibadan: Cambridge University Press, in association with the West African Languages Survey and the Institute of African Studies.
Sauvageot, Serge. 1965. *Description synchronique d'un dialecte wolof: le parler du dyolof*. Mémoires de l'Institut Fondamental d'Afrique Noire 73. Dakar: Institut Français de l'Afrique Noire.
Scancarelli, Janine. 1987. 'Grammatical Relations and Verb Agreement in Cherokee'. University of California at Los Angeles Ph.D. Dissertation.
Schachter, Paul. 1976. 'The Subject in Philippine Languages: Topic, Actor, Actor-topic, or None of the Above'. In Charles Li, ed., *Subject and Topic*, pp. 491–518. New York: Academic Press.
—— and Fé T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley, CA: University of California Press. [Reprinted 1983.]
Schaub, Willi. 1985. *Babungo*. Croom Helm Descriptive Grammars. London: Croom Helm.
Schiffman, Harold F. 1983. *A Reference Grammar of Spoken Kannada*. Seattle: University of Washington Press.
Schmid, Maureen Alicia. 1980. 'Co-occurrence Restrictions in Negative, Interrogative, and Conditional Clauses: A Cross-Linguistic Study.' State University of New York at Buffalo Ph.D. Dissertation.
Schmidt, Wilhelm. 1926. *Die Sprachfamilien und Sprachenkreise der Erde*. Heidelberg: Carl Winter.
Schuh, Russell G. 1998. *A Grammar of Miya*. University of California Publications in Linguistics 130. Berkeley, CA: University of California Press.
Schultze-Berndt, Eva. 2000. 'Simple and Complex Verbs in Jaminjung: A Study of Event Categorization in an Australian Language'. Katholieke Universiteit Nijmegen Ph.D. Dissertation.
Schwartz, Linda and Timothy Dunnigan. 1986. 'Pronouns and Pronominal Categories in Southwestern Ojibwe'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 285–322. Tübingen: Gunter Narr.
Schwarz, Oxana, Cornelia Stroh, and Aina Urdze. 2001. *BIB-KOM-TYP: Die Bibliographie zur Komitativ-Typologie*. München: Lincom Europa.
Scorza, David. 1985. *A Sketch of Au Morphology and Syntax*. Papers in New Guinea Linguistics 22. Pacific Linguistics, Series A-63. Canberra: Australian National University.
Scott, Graham. 1978. *The Fore Language of Papua New Guinea*. Pacific Linguistics, Series B-47. Canberra: Australian National University.
Sebeok, Thomas A. and Frances J. Ingemann. 1961. *An Eastern Cheremis Manual: Phonology, Grammar, Texts, and Glossary*. Indiana University Publications, Uralic and Altaic Series 5, Studies in Cheremis 9. Bloomington and The Hague: Indiana University and Mouton.
Sedlak, Philip A. S. 1975. 'Direct/Indirect Object Word Order: a Cross-Linguistic Analysis'. *Working Papers on Language Universals* 18: 117–64. Stanford: Stanford University Press.
Seiler, Hansjakob. 1983. *Possession as an Operational Dimension of Language*. Tübingen: Gunter Narr.
Seiler, Walter. 1985. *Imonda, a Papuan Language*. Pacific Linguistics, Series B-93. Canberra: Australian National University.
Seiler, Wolf. 1978. 'The Modalis Case in Inupiat'. *Working Papers of the Summer Institute of Linguistics (University of North Dakota)* 22: 71–85.
Seiter, William J. 1980. *Studies in Niuean Syntax*. New York: Garland Publishing Company.
Semiun, Agustinus. 1993. 'The Basic Grammar of Manggarai'. La Trobe University M.A. Thesis.
Senft, Gunter. 1986. *Kilivila, the Language of the Trobriand Islanders*. Berlin: Mouton de Gruyter.
Shell, Olive. 1952. 'Grammatical Outline of Krahô (Ge Family)'. *International Journal of American Linguistics* 18.3: 115–29.
Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.
—— 1998. 'Honorifics'. In Jacob L. Mey, ed., *Concise Encyclopedia of Pragmatics*, pp. 341–50. Amsterdam: Elsevier.
Shkarban, Lina I. and Gennadij E. Rachkov. (forthcoming.) 'Reciprocal, Sociative and Comitative Constructions in Tagalog'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Showalter, Catherine. 1986. 'Pronouns in Lyele'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 205–16. Tübingen: Gunter Narr.
Siewierska, Anna. 1984. *The Passive: A Contrastive Linguistic Analysis*. London: Croom Helm.
—— 1991. *Functional Grammar*. London and New York: Routledge.
Siewierska, Anna. 2004. 'On the Discourse Basis of Person Agreement'. In Tuija Virtanen, ed., *Discourse Approaches to Anaphora*. Berlin: Mouton de Gruyter.
Silverstein, Michael. 1972. 'Chinook Jargon: Language Contact and the Problem of Multi-level Generative Systems, II'. *Language* 48: 596–625.
—— 1976. 'Hierarchy of Features and Ergativity'. In R. M. W. Dixon, ed., *Grammatical Categories in Australian Languages*, pp. 112–71. Canberra: Australian National University.
Simon, Horst J. 1997. 'Die Diachronie der deutschen Anredepronomina aus der Sicht der Universalienforschung'. *Sprachtypologie und Universalienforschung* 50.3: 267–81.
—— 1999. 'Für eine grammatische Kategorie "Respekt" im Deutschen. Synchronie, Diachronie und Typologie der deutschen Anredepronomina'. Humboldt University Berlin Ph.D. Dissertation.
Sischo, William R. 1977. 'Michoacán Nahuatl'. In Ronald W. Langacker, ed., *Modern Aztec Grammatical Sketches*, Studies in Uto-Aztecan Grammar 2, pp. 307–80. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
Sjoberg, Andrée F. 1963. *Uzbek Structural Grammar*. Bloomington: Indiana University.
Skinner, Margaret Gardner. 1979. 'Aspects of Pa'anci Grammar'. University of Wisconsin Ph.D. Dissertation.
Skorik, P. Ja. 1961. *Grammatika čukotskogo jazyka*. Čast' pervaja. Moscow: Izdatel'stvo Akademii Nauk SSSR.
—— 1986. 'Kategorii imeni suščestvitel'nogo v čukotsko-kamčatskix jazykax'. In P. J. Skorik, ed., *Paleoaziatskie jazyki*, pp. 76–111. Leningrad: Nauka.
Smeets, Ineke. 1989. 'A Mapuche Grammar'. Rijksuniversiteit te Leiden Ph.D. Dissertation.
Smith-Stark, Thomas Cedric. 1974. 'The Plurality Split'. *Chicago Linguistic Society* 10: 657–61.
—— 1978. 'The Mayan Antipassive: Some Facts and Fictions'. In Nora England, Colette Craig, and Loanna Furbee, eds., *Papers in Mayan Linguistics*, pp. 169–87. Columbia: University of Missouri Press.
Sneddon, James N. 1975. *Tondano Phonology and Grammar*. Canberra: Australian National University.
—— 1996. *Indonesian: A Comprehensive Grammar*. New York: Routledge.
Sohn, Ho-min. 1975. *Woleaian Reference Grammar*. Honolulu: University of Hawaii Press.
—— 1994. *Korean*. Descriptive Grammars. London and New York: Routledge.
Solnit, David B. 1997. *Eastern Kayah Li: Grammar, Texts, Glossary*. Honolulu: University of Hawaii Press.
Song, Jae Jung. 1996. *Causatives and Causation: A Universal-Typological Perspective*. London and New York: Addison Wesley Longman.
—— 2001. *Linguistic Typology: Morphology and Syntax*. Harlow: Pearson Education.
Spagnolo, L. M. 1933. *Bari Grammar*. Verona: Missioni Africane.

Speiser, E. A. 1941. *Introduction to Hurrian*. The Annual of the American Schools of Oriental Research XX. New Haven: American Schools of Oriental Research.

Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.

Sridhar, S. N. 1990 *Kannada*. Descriptive Grammars Series. London and New York: Routledge.

Stairs, Emily F. and Elena E. de Hollenbach. 1981. 'Gramática huave'. In Glenn A. S. Kreger and Emily F. Stairs, eds., *Diccionario huave de San Mateo del Mar*, pp. 283–391. México, DF: Instituto Lingüístico de Verano.

Staley, William E. 1995. 'Referent Management in Olo: A Cognitive Perspective'. University of Oregon Ph.D. Dissertation.

Stassen, Leon. 1984. 'The Comparative Compared'. *Journal of Semantics* 3: 143–82.

—— 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

—— 1997. *Intransitive Predication*. Oxford: Clarendon Press.

—— 2000. 'AND-Languages and WITH-Languages'. *Linguistic Typology* 4.1: 1–54.

—— 2001. 'Predicative Possession'. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, eds., *Language Typology and Language Universals, An International Handbook 2*, pp. 954–60. Berlin and New York: Mouton de Gruyter.

Steele, Susan. 1977. 'On Being Possessed'. *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistic Society* 3: 114–31.

Stein, Mark J. 1981. 'Quantification in Thai'. University of Massachusetts Ph.D. Dissertation.

Stevens, K. N. 1972. 'The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data'. In E. E. David and P. B. Denes, eds., *Human Communication: A Unified View*, pp. 51–6. New York: McGraw-Hill.

Stevenson, R. C. 1969. *Bagirmi Grammar*. Linguistic Monograph Series 3. Khartoum: Sudan Research Unit, University of Khartoum.

Stolz, Thomas. 1996. 'Some Instruments Are Really Good Companions—Some Are Not. On Syncretism and the Typology of Instrumentals and Comitatives'. *Theoretical Linguistics* 23.1/2: 113–200.

—— 2001a. 'To be with X is to have X: Comitatives, Instrumentals, Locative, and Predicative Possession'. *Linguistics* 39.2: 321–50.

—— 2001b. 'Comitatives vs. Instrumentals vs. Agents'. In Walter Bisang, ed., *Aspects of Typology and Universals*, pp. 153–74. Berlin: Akademie Verlag.

—— 2001c. 'On Circum-Baltic Instrumentals and Comitatives'. In Östen Dahl and Maria Koptjevskaja-Tamm, eds., *Circum-Baltic Languages*, pp. 591–612. Amsterdam: John Benjamins.

—— 2001d. 'ORDINALIA—Linguistisches Neuland. Ein Typologenblick auf die Beziehung zwischen Kardinalia und Ordinalia und die Sonderstellung von EINS und ERSTER'. In Birgit Igla and Thomas Stolz, eds., *Was Ich Noch Sagen Wollt . . . A Multilingual Festschrift for Norbert Boretzky on Occasion of his 65th Birthday*, pp. 507–30. Berlin: Akademie Verlag.

—— and Cornelia Stroh. 2001. 'Wenn Komitative Instrumentale sind und umgekehrt'. In Winfried Boeder and Gerd Hentschel, eds., *Variierende Markierung von Nominalgruppen in Sprachen unterschiedlichen Typs*, pp. 387–411. Oldenburg: BIS Verlag.

Street, John C. 1963. *Khalkha Structure*. Indiana University Publications, Uralic and Altaic Series 24. Bloomington: Indiana University Press.

Strom, Clay. 1992. *Retuarã Syntax*. Studies in the Languages of Colombia 3. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Sulkala, Helena and Merja Karjalainen. 1992. *Finnish*. London and New York: Routledge.

Sun, Hongkai and Liu Lu. 1986. *Nuzu yuyan jianzhi (nusuyu)*. [A brief description of the language of the Nu nationality (Nusu language)]. Beijing: Minzu Chubanshe.

Sutton-Spence, Rachel and Bencie Woll. 1999. *The Linguistics of British Sign Language: An Introduction*. Cambridge: Cambridge University Press.

Swanton, John R. 1911. 'Haida'. In Franz Boas, ed., *Handbook of American Indian Languages 1*, Bureau of American Ethnology Bulletin 40, pp. 205–82. Washington, DC: Smithsonian Institution.

T

Tamura, Suzuko. 1988. *The Ainu Language*. Tokyo: Sanseido. [Reprinted in 2000, translated from Japanese into English by Sanseido Co. Ltd.]

—— 2000. *The Ainu Language*. ICHEL Linguistic Studies 2. Tokyo: Sanseido.

Tarpent, Marie-Lucie. 1987. 'A Grammar of the Nisgha Language'. University of Victoria (BC) Ph.D. Dissertation.

Tayeng, Aduk. 1976. *Milang Phrase-Book*. Shillong: Director of Information and Public Relations, Arunachal Pradesh.

Taylor, Allan R. 1996. 'Nonspeech Communication Systems'. In Ives Goddard, ed., *Handbook of North American Indians*, volume 17: Languages, pp. 275–89. Washington: Smithsonian Institution.

Taylor, Charles. 1985. *Nkore-Kiga*. Croom Helm Descriptive Grammars. London: Croom Helm.

Tchekhoff, Claude. 1987. '"Antipassif": Aspect imperfectif et autonomie du sujet'. *Bulletin de la Société de Linguistique de Paris* 82: 43–67.

Tenišev, È. R. et al., eds. 1997. *Jazyki mira: tjurkskie jazyki*. Moscow: Indrik.

Tenny, Carol. 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Studies in Linguistics and Philosophy 52. Dordrecht: Kluwer.

Tereščenko, N. M. 1966a. 'Vvedenie (samodijskie jazyki)'. In V. V. Vinogradov, ed., *Jazyki narodov SSSR, Volume 3: Finno-ugorskie i samodijskie jazyki*, pp. 363–75. Moscow and Leningrad: Nauka.

—— 1966b. 'Nganasanskij jazyk'. In V. V. Vinogradov, ed., *Jazyki Narodov SSSR, Volume 3: Finno-ugorskie i samodijskie jazyki,* pp. 416–37. Moscow and Leningrad: Nauka.

Terrill, Angela. 1999. 'A Grammar of Lavukaleve: A Papuan Language of the Solomon Islands'. Australian National University, Canberra, Ph.D. Dissertation.

—— 2003. *A Grammar of Lavukaleve*. Mouton Grammar Library 30. Berlin: Mouton de Gruyter.

Thelwall, Robin. 1981. 'Lexicostatistical Subgrouping and Lexical Reconstruction of the Daju Group'. In T. Schadeberg and L. Bender, eds., *Nilo-Saharan: Proceedings of the First Nilo-Saharan Linguistics Colloquium, Leiden, September 8–10, 1980*, pp. 167–85. Dordrecht: Foris.

Thomas, David D. 1955. 'Three Analyses of the Ilocano Pronoun System'. *Word* 11.2: 204–8.

Thompson, Laurence C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.

Thompson, Laurence C., M. Terry Thompson and Barbara S. Efrat. 1974. 'Some Phonological Developments in Straits Salish'. *International Journal of American Linguistics* 40.3: 182–96.

Thráinsson, Höskuldur. 1994. 'Icelandic'. In Ekkehard König and Johan van der Auwera, eds., *The Germanic Languages*, pp. 142–89. London and New York: Routledge.

Thurston, William R. 1982. *A Comparative Study in Anem and Lusi*. Pacific Linguistics, Series B-83. Canberra: Australian National University.

Tindale, Norman B. 1994. *Tribal Boundaries in Aboriginal Australia*. Berkeley: University of California Press. Map.

Tomlin, Russell. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.

Tompa, József. 1968. *Ungarische Grammatik*. Janua Linguarum, Series Practica 96. The Hague: Mouton.

Tonelli, Antonio. 1926. *Grammatica e glossario della lingua degli onaselknam della Terra del Fuoco*. Torino: Società Editrice Internazionale.

Topping, Donald M. (with the assistance of Bernadita C. Dungca). 1973. *Chamorro Reference Grammar*. Honolulu: University Press of Hawaii.

—— ,—— 1980. *Chamorro Reference Grammar*. Honolulu: University Press of Hawaii. [1st edn. 1973.]

Tosco, Mauro. 1991. *A Grammatical Sketch of Dahalo*. Hamburg: Helmut Buske Verlag.

Trail, Ronald L. 1970. *The Grammar of Lamani*. Summer Institute of Linguistics Publications in Linguistics and Related Fields 24. Norman: University of Oklahoma.

Traill, Anthony. 1994. *A !Xóõ Dictionary*. Köln: Köppe Verlag.

Traugott, Elizabeth Closs. 1989. 'On the Rise of Epistemic Meaning in English: an Example of Subjectification in Semantic Change'. *Language* 65: 31–55.

Trenga, G. 1947. *Le Bura-Mabang du Ouadaï*. Paris: Institut d'Ethnologie, Université de Paris.

Tryon, Darrell T. 1968. *Iai Grammar*. Pacific Linguistics, Series B-8. Canberra: Australian National University.

—— 1970. *An Introduction to Maranungku*. Pacific Linguistics, Series B-15. Canberra: Australian National University.

—— 1994. 'Language Contact-induced Language Change in the Eastern Outer Islands, Solomon Islands'. In Tom Dutton and Darrell T. Tryon, eds., *Language Contact and Change in the Austronesian World*, pp. 611–48. Berlin: Mouton de Gruyter.

Tsunoda, Tasaku. 1981. 'Split Case-marking Patterns in Verb-types and Tense/Aspect/Mood'. *Linguistics* 19: 389–438.

—— 1988. 'Antipassives in Warrungu and Other Australian Languages'. In Masayoshi Shibatani, ed., *Passive and Voice*, pp. 595–649. Amsterdam: John Benjamins.

Tsunoda, Tasaku. 1998. 'Applicative Constructions in Warrungu of Australia'. In Anna Siewierska and Jae Jung Song, eds., *Case, Typology and Grammar: In Honor of Barry J. Blake*, pp. 343–73. Amsterdam: John Benjamins.
—— (forthcoming a.) 'Reciprocal Constructions in Warrungu'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— (forthcoming b.) 'Reciprocal-reflexive Constructions in Djaru'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— Sumie Ueda, and Yoshiaki Itoh. 1995. 'Adpositions in Word-order Typology'. *Linguistics* 33: 741–61.
Tucker, A. N. and M. A. Bryan. 1966. *Linguistic Analyses: the Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.
—— and J. Tompo Ole Mpaayei. 1955. *A Maasai Grammar with Vocabulary*. London: Longmans, Green & Co.
Tuggy, David H. 1979. 'Tetelcingo Nahuatl'. In Ronald W. Langacker, ed., *Modern Aztec Grammatical Sketches*, Studies in Uto-Aztecan Grammar 2, pp. 1–140. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

U

Ubrjatova, E. I. 1985. *Jazyk noril'skix dolgan*. Novosibirsk: Nauka.
Učida, Norihiko. 1970. *Der Bengali-Dialekt von Chittagong: Grammatik, Texte, Wörterbuch*. Wiesbaden: Harrassowitz.
Ultan, Russell. 1972. 'Some Features of Basic Comparative Constructions'. *Working Papers on Language Universals (Stanford)* 9: 117–62.
—— 1978. 'The Nature of Future Tenses'. In Joseph H. Greenberg, ed., *Universals of Human Language*, pp. 83–123. Stanford: Stanford University Press.
Underhill, Robert. 1976. *Turkish Grammar*. Cambridge, MA: MIT Press.
Unseth, Peter. 1989. 'Sketch of Majang Syntax'. In M. Lionel Bender, ed., *Topics in Nilo-Saharan Linguistics*, pp. 97–127. Hamburg: Helmut Buske Verlag.

V

Vago, Robert M, ed. 1980. *Issues in Vowel Harmony*. Amsterdam: John Benjamins.
Valenzuela, Pilar M. 1997. 'Basic Verb Types and Argument Structures in Shipibo-Conibo'. University of Oregon M.A. Thesis.
—— 2000. 'Ergatividad escindida en Wariapano, Yaminawa y Shipibo-Konibo'. In Hein van der Voort and Simon van de Kerke, eds., *Indigenous Languages of Lowland South America*, pp. 111–128. Leiden: Research School of Asian, African, and Amerindian Languages (CNWS), University of Leiden.
—— 2002. 'Causativization and Transitivity in Shipibo-Konibo'. In Masayoshi Shibatani, ed., *The Grammar of Causation and Interpersonal Manipulation*, Typological Studies in Language 48, pp. 417–83. Amsterdam: John Benjamins.
—— 2003. 'Transitivity in Shipibo-Konibo Grammar'. University of Oregon Ph.D. Dissertation.
Valkama, Kari. 2000. *Grammatical Relations in Cebuano*. Helsinki: University of Helsinki, Department of General Linguistics.
van de Vijver, R. 1998. *The Iambic Issue: Iambs as a Result of Constraint Interaction*. HIL Dissertation 37. The Hague: Holland Academic Graphics.
van den Berg, Helma. 1995. *A Grammar of Hunzib*. Lincom Studies in Caucasian Linguistics 1. München: Lincom Europa.
van den Berg, Helma. 2001. *Dargi Folktales with Interlinear Translation and an Introduction to the Grammar*. Leiden: Research School of Asian, African and Amerindian Studies.
van den Berg, René. 1989. *A Grammar of the Muna Language*. Dordrecht: Foris.
van der Auwera, Johan and Vladimir Plungian. 1998. 'Modality's Semantic Map'. *Linguistic Typology* 2: 79–124.
van der Hulst, Harry. 1984. *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.
—— 1996. 'Separating Primary Accent and Secondary Accent'. In R. Goedemans, H. van der Hulst, and E. Visch, eds., *Stress Patterns of the World, Volume 1: Background*, HIL Publications 2, pp. 1–26. The Hague: Holland Academic Graphics.
—— 1999a. 'Word Accent'. In H. van der Hulst, ed., *Word Prosodic Systems in the Languages of Europe*, pp. 3–116. Berlin and New York: Mouton de Gruyter.
—— 1999b. 'Issues in Foot Typology'. In M. Davenport and S. J. Hannahs, eds., *Issues in Phonological Structure*, pp. 95–127. Amsterdam: John Benjamins.
van Driem, George. 1987. *A Grammar of Limbu*. Mouton Grammar Library 4. Berlin: Mouton de Gruyter.
—— 1991. *A Grammar of Dumi*. Berlin: Mouton de Gruyter.
van Eijk, Jan. 1997. *The Lillooet Language: Phonology, Morphology, Syntax*. Vancouver: University of British Columbia Press.
van Enk, Gerrit J. and Lourens de Vries. 1997 *The Korowai of Irian Jaya: Their Language in its Cultural Context*. Oxford: Oxford University Press.
van Klinken, Catharina Lumien. 1999. *A Grammar of the Fehan Dialect of Tetun, an Austronesian Language of West Timor*. Pacific Linguistics, Series C-155. Canberra: Australian National University.
van Oostendorp, Marc. 2000. *Phonological Projection: A Theory of Feature Content and Prosodic Structure*. Berlin and New York: Mouton de Gruyter.
van Staden, Miriam. 2000. 'Tidore: A Linguistic Description of a Language of the North Moluccas'. Leiden University Ph.D. Thesis.
Van Valin, Robert D. 1991. 'Semantic Parameters of Split Intransitivity'. *Language* 66: 221–60.
—— and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.
Vandame, Charles. 1968. *Grammaire kenga*. Paris: Afrique et Langage.
Vaxtin, N. B. 1997. 'Grenlandskij jazyk'. In A. P. Volodin et al., eds., *Jazyki mira: paleoaziatskie jazyki*, pp. 93–101. Moscow: Indrik.
Vendler, Zeno. 1967. 'Facts and Events'. In Zeno Vendler, *Linguistics in Philosophy*, pp. 12–146. Ithaca: Cornell University Press.
Verkuyl, Hendrik Jacob. 1971. *On the Compositional Nature of the Aspects*. Amsterdam: Reidel.
Veselinova, Ljuba. 1997. 'Suppletion in the Derivation of Ordinal Numerals: a Case Study'. In Benjamin Bruening, ed., *Proceedings of the 8th Student Conference in Linguistics*, pp. 429–47. Cambridge, MA: MIT Working Papers in Linguistics.
—— 2003. 'Suppletion in Verb Paradigms: Bits and Pieces of a Puzzle'. Stockholm University Ph.D. Dissertation.
Viñas Urquiza, María Teresa. 1974. *Lengua mataca*. Buenos Aires: Centro de Estudios Lingüísticos.
Vitale, Anthony J. 1981. *Swahili Syntax*. Dordrecht: Foris.
Vogt, Hans. 1971. *Grammaire de la langue géorgienne*. Oslo: Institute for Comparative Research in Human Culture.
Volodin, Aleksandr P. 1976. *Itel'menskij jazyk*. Leningrad: Nauka.
—— 1997. 'Itel'menskij jazyk'. In A. P. Volodin et al., eds., *Jazyki mira: paleoaziatskie jazyki*, pp. 60–72. Moscow: Indrik.
—— 2001. 'Imperative in Kerek'. In Viktor S. Xrakovskij, ed., *Typology of Imperative Constructions*, pp. 145–58. München: Lincom Europa.
—— (forthcoming.) 'Reciprocal Constructions in Itelmen (Kamchadal)'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— and P. Skorik. 1997. 'Čukotksij jazyk'. In A. P. Volodin et al., eds., *Jazyki mira: paleoaziatskie jazyki*, pp. 22–39. Moscow: Indrik.
Voorhoeve, Clemens L. 1965a. 'The Flamingo Bay Dialect of the Asmat Language'. University of Leiden Ph.D. Dissertation.
—— 1965b. *The Flamingo Bay Dialect of the Asmat Language*. Verhandelingen van het Koninklijk Instituut voor Taal, Land en Volkenkunde 46. The Hague: M. Nijhoff.
—— 1971. 'Miscellaneous Notes on Languages in West Irian, New Guinea'. In T. Dutton et al., eds., *Papers in New Guinea Linguistics* 14, pp. 47–114. Pacific Linguistics, Series A-28. Canberra: Australian National University.
Vorbichler, Anton. 1971. *Die Sprache der Mamvu*. Glückstadt: Verlag J. J. Augustin.
Vydrine, Valentin. (forthcoming.) 'Reciprocal and Sociative Constructions in Bamana'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.

W

Wade, Terence L. B. 1992. *A Comprehensive Russian Grammar*. Oxford: Blackwell. [reprinted in 1995].
Wald, Benji. 1975. 'Animate Concord in Northeast Coastal Bantu: its Linguistic and Social Implications as a Case of Grammatical Convergence'. *Studies in African Linguistics* 6: 267–314.
Walker, Dale F. 1976. *Grammar of the Lampung Language*. Jakarta: NUSA.
Walker, Rachel. 1999. 'Guaraní Voiceless Stops in Oral versus Nasal Contexts: An Acoustical Study'. *Journal of the International Phonetic Association* 29.1: 63–94.
Walsh, Michael James. 1976. 'The Murinypata Language of North-West Australia'. Australian National University Ph.D. Dissertation.
Warotamasikkhadhit, Udom. 1972. *Thai Syntax*. The Hague: Mouton.

Wash, Suzanne. 2001. 'Adverbial Clauses in Barbareño Chumash Narrative Discourse'. University of California, Santa Barbara Ph.D. Dissertation.
Watkins, Laurel. 1984. *A Grammar of Kiowa*. Lincoln: University of Nebraska Press.
Watters, James K. 1988. 'Topics in Tepehua Grammar'. University of California at Berkeley Ph.D. Dissertation.
Webb, Rebecca and Ted Supalla. 1994. 'Negation in International Sign'. In Inger Ahlgren, Brita Bergman, and Mary Brennan, eds., *Perspectives on Sign Language Structure. Papers from the Fifth International Symposium on Sign Language Research 1*, pp. 173–85. Durham: International Sign Linguistics Association and Deaf Studies Research Unit, University of Durham.
Weir, E. M. Helen. 1994. 'Nadëb'. In Peter Kahrel and René van den Berg, eds., *Typological Studies in Negation*, pp. 291–323. Amsterdam: John Benjamins.
Weiß, Helmut. 1999. 'Duplex negatio non semper affirmat: A Theory of Double Negation in Bavarian'. *Linguistics* 37.5: 819–46.
Welmers, William E. 1973. *African Language Structures*. Berkeley: University of California Press.
—— 1976. *A Grammar of Vai*. University of California Publications in Linguistics 84. Berkeley, CA: University of California Press.
Wendt, Heinz F. 1972. *Langenscheidts Praktisches Lehrbuch Türkisch*. Berlin: Langenscheidt.
Werner, Heinrich. 1997. *Die ketische Sprache*. Tungusco-Sibirica 3. Wiesbaden: Harrassowitz.
Wetzer, Harrie. 1996. *Nouniness and Verbiness: a Typological Study of Adjectival Predication*. Berlin: Mouton De Gruyter.
Whiteley, W. H. 1958. *A Short Description of Item Categories in Iraqw*. Kampala: East African Institute of Social Research.
Wiemer, Björn. (forthcoming.) 'Reciprocal and Reflexive Constructions in Polish'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
—— and Vladimir P. Nedjalkov. (forthcoming.) 'Reciprocal and Reflexive Constructions in German'. In Nedjalkov et al., eds., *Typology of Reciprocal Constructions*. München: Lincom Europa.
Wiering, Elisabeth. 1994. 'Some Major Structures of Doyayo Syntax'. In Elisabeth Wiering and Marinus Wiering, eds., *The Doyayo Language*, pp. 117–265. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
Wiesemann, Ursula. 1986a. 'The Pronoun Systems of some Jê and Macro-Jê Languages'. In Ursula Wiesemann, ed., *Pronominal Systems*, pp. 359–80. Tübingen: Gunter Narr.
Wiesemann, Ursula, ed. 1986b. *Pronominal Systems*. Tübingen: Gunter Narr.
Wilkins, David P. 1989. 'Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar'. Australian National University Ph.D. Thesis.
Willett, Thomas L. 1991. *A Reference Grammar of Southeastern Tepehuan*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
Williams, Stephen J. 1980. *A Welsh Grammar*. Cardiff: University of Wales Press.
Williamson, Kay. 1965. *A Grammar of the Kolokuma Dialect of Ijo*. West African Language Monographs 2. Cambridge: Cambridge University Press.
Willms, Alfred. 1972. *Grammatik der südlichen Beraberdialekte (Südmarokko)*. Afrikanistische Forschungen 6. Hamburg: J. J. Augustin.
Wilson, Darryl. 1974. *Suena Grammar*. Workpapers in Papua New Guinea Languages 8. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.
Wilson, Patrizia R. 1980. *Ambulas Grammar*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.
Wilson, Peter. 1970. *Simplified Swahili*. London: Longman.
Wise, Mary Ruth. 1990. 'Valence-changing Affixes in Maipuran Arawakan Languages'. In Doris Payne, ed., *Amazonian Linguistics, Studies in Lowland South American Languages*, pp. 89–116. Austin: University of Texas Press.
—— 1999. 'Small Language Families and Isolates in Peru'. In R. M. W. Dixon and Alexandra Aikhenvald, eds., *The Amazonian Languages*, pp. 307–40. Cambridge: Cambridge University Press.
Witkowski, Stanley R. and Cecil H. Brown. 1985. Climate, Clothing, and Body-part Nomenclature. *Ethnology* 24: 197–214.
Witucki, Jeannette. 1974. 'The Vowels of Chamorro'. In *Papers in Philippine Linguistics 6*, Pacific Linguistics, Series A-43, pp. 55–73. Canberra: Australian National University.
Wivell, Richard. 1981. 'Kairiru Grammar'. University of Auckland M.A. Thesis.
Wolfart, H. Christoph. 1973. *Plains Cree: A Grammatical Study*. Transactions of the American Philosophical Society, New Series 63, Part 5. Philadelphia: American Philosophical Society.
—— and Janet Carroll. 1981. *Meet Cree*. Edmonton: University of Alberta Press. [2nd edn.]
Wolff, H. Ekkehard. 1993. *Referenzgrammatik des Hausa: Zur Begleitung des Fremdsprachenunterrichts und zur Einführung in das Selbststudium*. Hamburg: LIT Verlag.
Wolff, John U. 1966. *Beginning Cebuano 1*. New Haven and London: Yale University Press.
Woolford, Ellen. 1993. 'Symmetric and Asymmetric Passives'. *Natural Language and Linguistic Theory* 11: 679–728.
Woollams, Geoffrey. 1996. *A Grammar of Karo Batak, Sumatra*. Pacific Linguistics, Series C-130. Canberra: Australian National University.
Wurm, Stephen A. 1982. *Papuan Languages of Oceania*. Tübingen: Gunter Narr.
Wurm, Stephen A., ed. 1975. *Papuan languages and the New Guinea linguistic scene. New Guinea area languages and language study, Volume 1*. Pacific Linguistics, Series C, No. 38. Canberra: Australian National University.
Wurzel, Wolfgang Ullrich. 1984. *Flexionsmorphologie und Natürlichkeit: Ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie-Verlag.

X

Xajdakov, S. M. 1963. 'Principy raspredelenija imen suščestvitel'nyx po grammatičeskim klassam v lakskom jazyke'. *Studia Caucasica* 1: 48–55.
Xrakovskij, Victor S., ed. 2001. *Typology of Imperative Constructions*. München: Lincom Europa. [Revised and expanded translation of Xrakovskij, Viktor S., ed. 1992. *Tipologija imperativnyx konstrukcij*. St. Petersburg: Nauka.]
Xu Lin and Yansun Zhao. 1984. *Baiyu jianzhi*. [A brief description of the Bai language]. Beijing: Minzu Chubanshe.
Xu Weiyan. 1994. 'A Study of Chinese Color Terminology'. Australian National University Ph.D. Dissertation.

Y

Young, Robert W. and William Morgan. 1987. *The Navajo Language: A Grammar and Colloquial Dictionary*. Albuquerque: University of New Mexico Press [revised edn.]
—— —— 1992. *Analytical Lexicon of Navajo*. Albuquerque: University of New Mexico Press.

Z

Zadeh, Lotfi A. 1996. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh*. Edited by G. J. Klir and B. Yuan. Singapore and River Edge, NJ: World Scientific.
Zaliznjak, A. A. 1964 'K voprosu o grammatičeskix kategorijax roda i oduševlennosti v sovremennom russkom jazyke'. *Voprosy Jazykoznanija* 4: 25–40.
Zanuttini, Raffaella. 1997. *Negation and Clausal Structure: a Comparative Study of Romance Languages*. Oxford: Oxford University Press.
Zaręba, Alfred. 1984–5. 'Osobliwa zmiana rodzaju naturalnego w dialektach polskich'. *Zbornik Matice Srpske za Filologiju i Lingvistiku* 17–18: 243–7.
Zec, D. 1988. 'Sonority Constraints on Prosodic Structure'. Stanford University Ph.D. Dissertation.
Zeshan, Ulrike. 2004a. 'Interrogative Constructions in Sign Languages—Cross-linguistic Perspectives'. *Language* 80.1: 7–39.
—— 2004b. 'Hand, Head and Face—Negative Constructions in Sign Language'. *Linguistic Typology* 8.1: 1–57.
Ziervogel, Dirk, Jacobus A. Louw and Petrus C. Taljaard. 1981. *A Handbook of the Zulu Language*. Pretoria: J. L. van Schaik.
Zipf, George. 1935. *The Psychobiology of Language: An Introduction to Dynamic Philology*. Cambridge, MA: MIT Press.
Žirkov, L. I. 1955. *Lakskij jazyk*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Language Name Index

This section includes a list of all *WALS* languages, identified by page numbers in the texts accompanying the maps if they are cited there and by the numeric code used in the Genealogical Language List. In addition to the *WALS* names for these languages are (i) names of families, subfamilies, and genera, shown in boldface; (ii) the names used in *Ethnologue* (Grimes 2000), when these are different from the *WALS* names; and (iii) other alternative names, either ones which deviate from the syntax of names in the Genealogical Language List, which places most modifiers in parentheses after the head (see §3.3 of the Introduction) (e.g. *Chalcatongo Mixtec* as an alternative for the *WALS* name *Mixtec (Chalcatongo)*), or ones that are alternative names that are used in some sources, primarily older sources (e.g. *Papago* for *O'odham*). Where there is a one-to-one correspondence between a *WALS* language and the *Ethnologue* equivalent, the equivalence is stated as in "Ahtena, Ethn. for Ahtna 87.1". Where there is not a one-to-one correspondence, either because a single *WALS* language corresponds to a number of *Ethnologue* languages or because a *WALS* language corresponds to a variety of a single *Ethnologue* language, the citation takes the form illustrated by "Aka-Bea, Ethn., see Aka-Biada 30.1".

A-Pucikwar 30.1
A-Pucikwar, Ethn., see Great Andamanese 30.1
Aari 7.5 p. 139
Abau 52.5
Abaza 19 pp. 442, 443
Abbéché Chad Arabic, see Arabic (Abbéché Chad) 7.6
Abenaki (Western) 89.1
Abidji 3.9
Abipón 199
Abkhaz 19 pp. 18, 43, 83, 190, 195, 199, 203, 247, 299, 400, 426, 427, 442, 450, 503, 506, 510
Abnaki (Western), Ethn. for Abenaki (Western) 89.1
Abo, Ethn. for Bankon 3.3.1
Aborlan Tagbanwa, see Tagbanwa (Aborlan) 31.6.3
Abu Arapesh, see Arapesh (Abu) 51.1
Abulas, Ethn. for Ambulas 52.1
Abun 32.4 pp. 374, 375, 466
Abzakh Adyghe, see Adyghe (Abzakh) 19
Aceh, Ethn. for Acehnese 31.6.10
Acehnese 31.6.10 p. 107
Achagua 174
Achang 26.2.4
Aché 176.6
Achí 132
Achí (Cubulco), Ethn., see Achí 132
Achí (Rabinal), Ethn., see Achí 132
Acholi 6.3.5
Achuar 162
Achuar-Shiwiar, Ethn. for Achuar 162
Achumawi 117.2
Acoma 106 pp. 35, 83, 102, 146, 238, 362, 438
Acooli, see Acholi 6.3.5
Adamawa-Ubangian 3.1 pp. 111, 283, 287
Adamorobe Sign Language 207 pp. 558, 561
Adi, Ethn., see Bokar 26.2.9
Adi, Ethn., see Bori 26.2.9
Adi, Ethn., see Gallong 26.2.9
Adi, Ethn., see Milang 26.2.9
Adioukrou 3.9
Adyghe (Abzakh) 19 pp. 151, 358
Adyghe (Shapsugh) 19 pp. 151, 358
Adyghe (Temirgoy) 19 pp. 151, 358
Adyghe, Ethn., see Adyghe (Abzakh) 19
Adyghe, Ethn., see Adyghe (Shapsugh) 19
Adyghe, Ethn., see Adyghe (Temirgoy) 19
Adynyamathanha 84.17
Adzera 31.3.1
Aekyom, Ethn. for Awin 64
Afade, Ethn. for Kotoko 7.2.1
Afar, Ethn. for Qafar 7.3.3
Afrikaans 9.5 p. 315
Afro-Asiatic 7 pp. 15, 27, 35, 122, 127, 135, 175, 194, 269, 283, 400, 415, 427, 435, 447
Agarabi 57.8
Agaw (Western), Ethn., see Kemant 7.3.2
Agaw, see Awngi 7.3.2
Agaw, see Bilin 7.3.2
Agaw, see Kemant 7.3.2
Aghem 3.3.1
Aghu 57.3 p. 374
Aghul 20.1.3 p. 299
Agni, see Anyi 3.9
Agta (Casiguran Dumagat), Ethn. for Dumagat (Casiguran) 31.6.4
Agta (Central Cagayan), Ethn. for Agta (Central) 31.6.4
Agta (Central) 31.6.4
Aguacatec 132 p. 67
Aguacateco, Ethn. for Aguacatec 132
Aguacatenango Tzeltal, see Tzeltal (Aguacatenango) 132
Aguaruna 162
Ahaggar Tuareg, see Tuareg (Ahaggar) 7.1
Ahtena, Ethn. for Ahtna 87.1
Ahtna 87.1 pp. 39, 523
Aikaná 174 p. 50
Aiki, see Runga 6.8
Ainu 16 pp. 58, 147, 163, 178, 187, 223, 235, 310, 326, 327a, 523
Ainu, Ethn. for Aynu 11.3
Air Tuareg, see Tuareg (Air) 7.1
Äiwo, see Ayiwo 83
Aizi 3.8
Aizi (Aproumu), Ethn., see Aizi 3.8
Aja 6.2.2
Ajia Varvara Romani, see Romani (Ajia Varvara) 9.7
Ajië 31.3.1
Aka 3.3.1
Aka-Bea, Ethn., see Aka-Biada 30.1
Aka-Biada 30.1 p. 43
Aka-Cari 30.1
Aka-Kede 30.1
Akan 3.9
Akawaio 175
Akha 26.2.4
Akhvakh 20.1.1
Akkadian p. 569
Aklan, see Aklanon 31.6.3
Aklanon 31.6.3 p. 66
Akposos, Ethn. for Kposo 3.9
Akwa 3.3.1
Ala'ala 31.3.1
Alabama 94 p. 115
Alabugat Tatar-Noghay, see Tatar-Noghay (Alabugat) 11.3
Alacalufan 203
Alacatlatzala Mixtec, see Mixtec (Alacatlatzala) 124.4
Alagwa 7.3.4
Alamblak 52.3 pp. 114, 322, 366
Alas 31.6.10
Alawa 84.13
Albanian 9.1 pp. 50, 58, 83, 250, 299, 315
Albanian (Gheg), Ethn., see Albanian 9.1
Albanian (Tosk), Ethn., see Albanian 9.1
Alekano, Ethn. for Gahuku 57.8
Alemannisch, Ethn., see Alsatian 9.5
Alemannisch, Ethn., see German (Appenzell) 9.5
Alemannisch, Ethn., see German (Ostschweiz) 9.5
Alemannisch, Ethn., see German (Thurgau) 9.5
Alemannisch, Ethn., see German (Zurich) 9.5
Aleut 86 pp. 78, 83, 147
Aleut (Eastern) 86
Aleut, Ethn., see Aleut (Eastern) 86
Algic 89 p. 35
Algonquian 89.1 pp. 87, 90, 115, 127, 131, 283, 339
Algonquin 89.1
Alladian 3.9
Allemannisch, Ethn., see German (Bern) 9.5
Allentiac 201
Alsatian 9.5
Alsea 114.1
Altai (Southern) 11.3
Altaic 11 pp. 50, 127, 194, 523
Alune 31.2 p. 235
Alutiiq 86
Alutor 14.1
Alyawarr, Ethn. for Alyawarra 84.17
Alyawarra 84.17 pp. 70, 526
Ama, Ethn. for Nyimang 6.3.7
Amadiya Neo-Aramaic, see Neo-Aramaic (Amadiya) 7.6
Amahuaca 168 pp. 83, 526
Amanab 46
Amara 31.3.1
Amarakaeri 171
Ambae (East), Ethn. for Ambae (Lolovoli Northeast) 31.3.1
Ambae (Lolovoli Northeast) 31.3.1
Ambai 31.3.2
Ambrym, see Lonwolwol 31.3.1
Ambulas 52.1 pp. 175, 178
Amdo 26.2.3
Amele 57.14 pp. 70, 83, 102, 142, 146, 230, 242, 243, 303, 418, 430, 446, 491, 495, 506
American Sign Language 207 pp. 558, 559, 560, 561, 565
Amharic 7.6 pp. 35, 127, 182, 422, 569
Ami, Ethn. for Emmi 84.3.6
Amis 31.4 p. 390
Amo 3.3.7
Ampeeli 57.1
Ampeeli-Wojokeso, Ethn. for Ampeeli 57.1
Amuesha 174 pp. 407, 523
Amuzgo 124.1
Amuzgo (Guerrero), Ethn., see Amuzgo 124.1
Amuzgoan 124.1
Anamgura, Ethn. for Anamuxra 57.14
Anamuxra 57.14
Ancash Quechua, see Quechua (Ancash) 169
/Anda, Ethn. for //Ani 1.1
Andamanese 30 p. 43
Andamanese, see A-Pucikwar 30.1
Andamanese, see Aka-Biada 30.1
Andamanese, see Aka-Cari 30.1
Andamanese, see Aka-Kede 30.1
Andamanese, see Great Andamanese 30.1
Andi 20.1.1
Andoke 147 pp. 18, 79
Andoque, Ethn. for Andoke 147
Aneityum, Ethn. for Anejom 31.3.1
Anejom 31.3.1 p. 414
Anem 76 p. 243
Angaatiha 57.1
Angami 26.2.7 p. 154
Angan 57.1
Angas 7.2.4
Anggor 47
Anglo Ewe, see Ewe (Anglo) 3.9
Angolar 206
Angor, Ethn. for Anggor 47
Anguthimri 84.17
//Ani 1.1 p. 82
Anindilyakwa 84.7.1 pp. 138, 183, 346

Annaberg 53.1
Anong 26.2.11
Anson Bay 84.3.1
Anuak, Ethn. for Anywa 6.3.5
Anufo 3.9
Anyi 3.9
Anyin, Ethn. for Anyi 3.9
Anyula, see Yanyuwa 84.17
Anywa 6.3.5 p. 282
Ao 26.2.7 p. 115
Ap Ma, Ethn. for Kambot 53.2
Apache (Chiricahua) 87.1 p. 523
Apache (Jicarilla) 87.1 pp. 522, 523
Apache (Mescalero-Chiricahua), Ethn. for Apache (Chiricahua) 87.1
Apache (Western) 87.1 pp. 513, 523, 526
Apalaí 175 pp. 291, 319
Apatani 26.2.9 p. 354
Apinayé 177.2
Apoze Kela, see Kela (Apoze) 31.3.1
Appenzell German, see German (Appenzell) 9.5
Apurinã 174 pp. 38, 46, 178, 287, 415, 502, 503
Arabana 84.17 p. 198
Arabela 158
Arabic (Abbéché Chad) 7.6
Arabic (Baharna Spoken), Ethn. for Arabic (Bahrain) 7.6
Arabic (Bahrain) 7.6
Arabic (Bani-Hassan) 7.6
Arabic (Beirut) 7.6
Arabic (Borno Nigerian) 7.6
Arabic (Cairene), see Arabic (Egyptian) 7.6
Arabic (Chadian Spoken) (= Arabic (Shuwa)), Ethn., see Arabic (Abbéché Chad) 7.6
Arabic (Cyproit Spoken), Ethn., see Arabic (Kormakiti) 7.6
Arabic (Eastern Libyan) 7.6
Arabic (Egyptian Spoken), Ethn. for Arabic (Egyptian) 7.6
Arabic (Egyptian) 7.6 pp. 43, 86, 122, 155, 311, 323, 362, 450, 502, 503
Arabic (Gulf Spoken), Ethn. for Arabic (Gulf) 7.6
Arabic (Gulf) 7.6 pp. 338, 359
Arabic (Hijazi Spoken), Ethn. for Arabic (Hijazi) 7.6
Arabic (Hijazi) 7.6
Arabic (Iraqi) 7.6 p. 362
Arabic (Kormakiti) 7.6
Arabic (Lebanese) 7.6
Arabic (Libyan Spoken), Ethn., see Arabic (Eastern Libyan) 7.6
Arabic (Mesopotamian Spoken), Ethn. for Arabic (Iraqi) 7.6
Arabic (Modern Standard) 7.6 pp. 10, 27, 83, 86, 127, 138, 139, 195, 223, 495
Arabic (Moroccan Spoken), Ethn. for Arabic (Moroccan) 7.6
Arabic (Moroccan) 7.6 p. 362
Arabic (Negev) 7.6
Arabic (North Levantine Spoken) 7.6
Arabic (North Levantine Spoken), Ethn., see Arabic (Beirut) 7.6
Arabic (North Levantine Spoken), Ethn., see Arabic (Lebanese) 7.6
Arabic (North Levantine Spoken), Ethn., see Arabic (Syrian) 7.6
Arabic (Palestinian) 7.6
Arabic (San'ani) 7.6 p. 572
Arabic (Sanaani), Ethn. for Arabic (San'ani) 7.6
Arabic (Shuwa) (= Arabic (Chadian Spoken)), Ethn., see Arabic (Borno Nigerian) 7.6
Arabic (South Levantine Spoken), Ethn. for Arabic (Palestinian) 7.6
Arabic (South Levantine Spoken), Ethn., see Arabic (Negev) 7.6
Arabic (Standard), Ethn. for Arabic (Modern Standard) 7.6
Arabic (Syrian) 7.6 pp. 330, 338, 375
Arabic (Tunisian Spoken), Ethn. for Arabic (Tunisian) 7.6
Arabic (Tunisian) 7.6
Arakanese (Marma) 26.2.4
Arakanese, Ethn. for Arakanese (Marma) 26.2.4
Araki 31.3.1
Aralle-Tabulahan 31.6.9
Aranda, see Arrernte 84.17
Arandai 59.2
Arandic p. 164
Araona 186 pp. 206, 399
Arapaho 89.1
Arapesh 51.1 pp. 115, 127, 135, 147
Arapesh (Abu) 51.1
Arapesh (Bumbita), Ethn., see Arapesh 51.1
Arára (Rondônia), Ethn. for Karó (Arára) 176.5
Arára Karó, see Karó (Arára) 176.5
Arauan 184
Araucanian 202
Araucanian, see Mapudungun 202
Arawak 174 pp. 142, 335, 386, 523
Arawakan 174 pp. 239, 291, 523, 548
Arbel Jewish Neo-Aramaic, see Neo-Aramaic (Arbel Jewish) 7.6
Arbore 7.3.3
Archi 20.1.3 pp. 30, 35, 66, 127, 234, 299, 415, 438, 439
'Are'are 31.3.1
Argentine Sign Language, Ethn. for Lengua de Señas Argentina 207
Aribwatsa 31.3.1
Arikara 103
Arikem 176.1
Arizona Tewa, see Tewa (Arizona) 105
Armenian 9.2
Armenian (Eastern) 9.2 pp. 35, 119, 127, 178, 268, 298
Armenian (Iranian) 9.2
Armenian (Western) 9.2
Armenian, Ethn., see Armenian (Eastern) 9.2
Armenian, Ethn., see Armenian (Iranian) 9.2
Armenian, Ethn., see Armenian (Western) 9.2
Arop-Lokep 31.3.1
Arosi 31.3.1
Arrarnta (Western), Ethn. for Arrernte (Western) 84.17
Arrarnta (Western), Ethn., see Arrernte 84.17
Arrernte 84.17 pp. 70, 71
Arrernte (Eastern), Ethn. for Arrernte (Mparntwe) 84.17
Arrernte (Eastern), Ethn., see Arrernte 84.17
Arrernte (Mparntwe) 84.17 pp. 235, 526
Arrernte (Western) 84.17
Aruak 135.1
Aruamu, Ethn. for Mikarew 53.3
Aruma, Ethn. for Wapishana 174
Asháninca, Ethn. for Campa (Axininca) 174
Asheninca Pajonal, Ethn. for Campa Pajonal Asheninca 174
Aslian 29.1.1 p. 573
Asmat 57.2 pp. 43, 102, 171, 178, 371, 486, 546, 547
Asmat-Kamoro 57.2
Assamese 9.7
Assangori, Ethn. for Sungor 6.3.9
Assyrian Neo-Aramaic, see Neo-Aramaic (Assyrian) 7.6
Assyrian, see Neo-Aramaic (Assyrian) 7.6
Asturian 9.9
Asu, Ethn. for Pare 3.3.1
Asuriní 176.6
Atacameño 194
Atakapa 97 p. 526
Atatlahuca Mixtec, see Mixtec (Atatlahuca) 124.4
Atayal 31.1 pp. 83, 282
Atayalic 31.1
Atchin 31.3.1
Ateso, see Teso 6.3.5
Athapaskan 87.1 pp. 115, 283, 523, 527, 569
Athpare 26.2.3
Athpariya, Ethn. for Athpare 26.2.3
Atikamekw 89.1 p. 526
Atlantic 3.2
Atoni, Ethn. for Timorese 31.2
Atsugewi 117.2
Atynyamatana, see Adynyamathanha 84.17
Au 51.3 p. 111
Auca, see Waorani 157
Aukan, Ethn. for Ndyuka 206
Aulua 31.3.1
Aunalei, Ethn. for One 51.4
Auslan 207 pp. 558, 559, 561, 564
Australian 84 pp. 19, 43, 67, 70, 78, 99, 194, 255, 318, 319, 338, 347, 438, 439, 442, 443, 487
Australian Sign Language, Ethn. for Auslan 207
Austro-Asiatic 29 pp. 35, 43, 83, 86, 131, 183, 269
Austronesian 31 pp. 43, 67, 87, 90, 94, 107, 111, 115, 127, 139, 162, 163, 165, 183, 194, 199, 223, 235, 259, 269, 283, 331, 335, 351, 355, 370, 395, 400, 401, 438, 442, 443, 487
Auyana 57.8
Avar 20.1.1 pp. 83, 474, 572
Avar-Andic-Tsezic 20.1.1
Avatime 3.9
Avokaya 6.2.6 p. 339
Awa 57.8
Awa Pit 140 pp. 342, 343, 367, 379, 415
Awa-Quaiquer, Ethn. for Awa Pit 140
Awadhi 9.7 p. 66
Awin 64
Awin-Pare 64
Awiyaana, Ethn. for Auyana 57.8
Awju-Dumut 57.3
Awngi 7.3.2
Awtuw 52.2 pp. 258, 434, 487
Awyi 46
Awyu (Miaro), Ethn. for Pisa 57.3
Axininca Campa, see Campa (Axininca) 174
Ayacucho Quechua, see Quechua (Ayacucho) 169
Ayamán, see Ayomán 149
Ayiwo 83
Aymara 170 pp. 15, 122, 123, 147, 298, 307, 400
Aymara (Central), Ethn., see Aymara 170
Aymara (Southern), Ethn., see Aymara 170
Aymaran 170
Aynu 11.3
Ayomán 149
Ayoreo 196
Ayt Seghrouchen Middle Atlas Berber, see Berber (Ayt Seghrouchen Middle Atlas) 7.1
Ayutla Mixtec, see Mixtec (Ayutla) 124.4
Azari (Iranian) 11.3
Azerbaijan Talysh, see Talysh (Azerbaijan) 9.8
Azerbaijani 11.3 p. 523
Azerbaijani (North), Ethn., see Azerbaijani 11.3
Azerbaijani (South), Ethn. for Azari (Iranian) 11.3
Azerbaijani (South), Ethn., see Azerbaijani 11.3
Aztecan 108.1
Baadi, Ethn. for Bardi 84.16
Baale 6.3.8
Babatana, Ethn., see Sisiqa 31.3.1
Babole 3.3.1
Babungo 3.3.1 pp. 259, 263, 307, 495, 497, 523
Bachajón Tzeltal, see Tzeltal (Bachajón) 132
Bachamal 84.3.1
Badaga 25.4
Bade 7.2.4
Bädi Kanum, see Kanum (Bädi) 63
Badimaya 84.17
Bafia 3.3.1
Bafut 3.3.1
Baga Sitemu 3.2.2
Bagandji, see Paakantyi 84.17
Bagirmi 6.2.1 pp. 43, 147
Bagiro 6.2.1
Bagri 9.7
Bagvalal 20.1.1 pp. 150, 298
Bahinemo 52.3
Bahnar (Plei Bong-Mang Yang) 29.1.2
Bahnar, Ethn., see Bahnar (Plei Bong-Mang Yang) 29.1.2
Bahnaric 29.1.2
Bahrain Arabic, see Arabic (Bahrain) 7.6
Bahuana 174
Bai 26.2.1 p. 390
Baining 78
Baining-Taulil 78
Baiso, Ethn. for Bayso 7.3.3
Bajau 31.6.6
Bajau (Indonesian), Ethn. for Bajau 31.6.6
Bajau (Semporna) 31.6.6
Baka (in Cameroon) 3.1 p. 142
Baka (in Sudan) 6.2.1
Baka, Ethn. for Baka (in Cameroon) 3.1
Baka, Ethn. for Baka (in Sudan) 6.2.1
Bakairí 175
Bakueri 3.3.1
Bakundu 3.3.1
Bakundu-Balue, Ethn. for Bakundu 3.3.1

Balangao 31.6.4
Balangingi Sama, see Sama (Balangingi) 31.6.6
Balanta 3.2.1
Balanta-Ganja, Ethn., see Balanta 3.2.1
Balanta-Kentohe, Ethn., see Balanta 3.2.1
Balantak 31.6.9
Balawaia, see Sinaugoro 31.3.1
Balese, see Lese 6.2.5
Bali Sign Language, Ethn. for Kata Kolok 207
Bali-Vitu 31.3.1
Bali, Ethn. for Balinese 31.6.10
Balinese 31.6.10
Balochi (Eastern), Ethn., see Baluchi 9.8
Balti 26.2.3
Baltic 9.3 pp. 58, 270
Baluchi 9.8
Balundu-Bima, Ethn. for Londo 3.3.1
Bambam 31.6.9
Bambara 3.10.2 pp. 295, 306, 366, 522
Bambili, Ethn. for Mbili 3.3.1
Bamun 3.3.1
Bana 7.2.1
Banawá 184 p. 70
Banda-Bambari, Ethn., see Linda 3.1
Banda-Linda, see Linda 3.1
Bandi 3.10.2
Bandjalang 84.17 pp. 267, 526
Bandjalang (Casino) 84.17
Bandjalang (Waalubal) 84.17 p. 526
Bandjalang (Yugumbir) 84.17 p. 526
Bandjalang, Ethn., see Bandjalang (Casino) 84.17
Bandjalang, Ethn., see Bandjalang (Waalubal) 84.17
Bandjalang, Ethn., see Bandjalang (Yugumbir) 84.17
Bandjalang, Ethn., see Gidabal 84.17
Banggai 31.6.9
Banggarla 84.17
Bangi, Ethn. for Bobangi 3.3.1
Bangla, see Bengali 9.7
Bani-Hassan Arabic, see Arabic (Bani-Hassan) 7.6
Baniata p. 183
Baniwa 174 p. 523
Bankon 3.3.1
Banoni 31.3.1
Bantik 31.6.9
Bantoid 3.3.1
Bantu languages pp. 11, 82, 98, 111, 122, 126, 131, 139, 211, 283, 291, 442, 443, 483, 573
Banz, see Wahgi 57.5
Baonan 11.1
Baoulé, Ethn. for Baule 3.9
Baoule, see Baule 3.9
Baraba Tatar, see Tatar (Baraba) 11.3
Baragaunle 26.2.3
Barai 57.12
Baram Kayan, see Kayan (Baram) 31.6.1
Barambu 3.1
Barasana, Ethn. for Barasano 145
Barasano 145 pp. 43, 46, 307, 315, 407, 419, 523
Barasano (Northern) 145 p. 523
Barbacoan 140
Barbareño Chumash, see Chumash (Barbareño) 123
Barbareño, Ethn. for Chumash (Barbareño) 123
Bardi 84.16
Baré 174 pp. 250, 291, 523
Bare'e, see Pamona 31.6.9
Barea, see Nara (in Ethiopia) 6.3.4
Bargam, Ethn. for Mugil 57.14
Bari 6.3.5 p. 414
Bari, Ethn., see Kukú 6.3.5
Bariba 3.5
Baric 26.2.2
Baruya 57.1
Barya, see Nara (in Ethiopia) 6.3.4
Basa, see Basaá 3.3.1
Basaá 3.3.1
Basaa, Ethn. for Basaá 3.3.1
Basaburua and Imoz Basque, see Basque (Basaburua and Imoz) 8
Basari 3.2.1
Bashkir 11.3 p. 298
Basque 8 pp. 43, 118, 147, 151, 178, 195, 203, 223, 250, 269, 359, 378, 399, 400, 447, 463, 491, 510, 511, 530
Basque (Basaburua and Imoz) 8
Basque (Bidasoa Valley) 8
Basque (Gernica) 8
Basque (Hondarribia) 8
Basque (Lekeitio) 8
Basque (Northern High Navarrese) 8
Basque (Oñati) 8
Basque (Roncalese) 8
Basque (Sakana) 8
Basque (Souletin) 8
Basque (Zeberio) 8
Basque, Ethn., see Basque (Basaburua and Imoz) 8
Basque, Ethn., see Basque (Bidasoa Valley) 8
Basque, Ethn., see Basque (Gernica) 8
Basque, Ethn., see Basque (Hondarribia) 8
Basque, Ethn., see Basque (Lekeitio) 8
Basque, Ethn., see Basque (Northern High Navarrese) 8
Basque, Ethn., see Basque (Oñati) 8
Basque, Ethn., see Basque (Roncalese) 8
Basque, Ethn., see Basque (Sakana) 8
Basque, Ethn., see Basque (Zeberio) 8
Batad Ifugao, see Ifugao (Batad) 31.6.4
Batak 31.6.3
Batak (Karo) 31.6.10 pp. 379, 400
Batak (Toba) 31.6.10
Batak Alas-Kluet, Ethn. for Alas 31.6.10
Bathari 7.6
Bats, Ethn. for Tsova-Tush 20.2
Batsbi, see Tsova-Tush 20.2
Bau 57.14
Baule 3.9
Baure 174
Bauzi 37
Bavarian German, see German (Bavarian) 9.5
Bavarian, Ethn. for German (Bavarian) 9.5
Bavarian, Ethn., see German (Upper Austrian) 9.5
Bavarian, Ethn., see German (Viennese) 9.5
Bawm 26.2.7 pp. 358, 427, 451
Bayali, Ethn., see Dharumbal 84.17
Bayso 7.3.3
Bedauye, see Beja 7.3.1
Bedawi, Ethn. for Beja 7.3.1
Beembe 3.3.1
Begak-Ida'an 31.6.1 p. 231
Behdini, Ethn., see Kurdish (Central) 9.8
Beirut Arabic, see Arabic (Beirut) 7.6
Beja 7.3.1 pp. 86, 110, 258, 315, 487, 572
Belgian Sign Language, Ethn., see Vlaamse Gebarentaal 207
Belhare 26.2.3 pp. 94, 98
Belhariye, Ethn. for Belhare 26.2.3
Bella Coola 110.1
Belorussian 9.10 p. 523
Bemba 3.3.1 p. 523
Bena-Lulua 3.3.1
Benabena 57.8
Bench, Ethn. for Gimira 7.5
Beng 3.10.1
Benga 3.3.1
Bengali 9.7
Bengali (Chittagong) 9.7 p. 206
Beni Amer Tigré, see Tigré (Beni Amer) 7.6
Benue-Congo 3.3 pp. 522, 523
Beothuk 90 p. 526
Bera, Ethn., see Bira 3.3.1
Berber 7.1 pp. 15, 83, 139, 211, 283, 331, 335, 569
Berber (Ayt Seghrouchen Middle Atlas) 7.1
Berber (Chaouia) 7.1
Berber (Figuig) 7.1
Berber (Mauritania), see Zenaga 7.1
Berber (Middle Atlas) 7.1 pp. 83, 86, 210, 400, 401
Berber (Mzab) 7.1
Berber (Rif) 7.1 pp. 182, 363
Berber (Siwa) 7.1
Berber (Wargla) 7.1
Berbice Dutch Creole 206
Berik 40
Berlin German, see German (Berlin) 9.5
Bern German, see German (Bern) 9.5
Berom, Ethn. for Birom 3.3.9
Berta 6.1 p. 83
Besisi, Ethn. for Mah Meri 29.1.1
Bessarabian Yiddish, see Yiddish (Bessarabian) 9.5
Betawi, Ethn. for Melayu Betawi 206
Bété 3.8 p. 14
Bété (Daloa), Ethn., see Bété 3.8
Betoi 151
Bezhta 20.1.1 pp. 438, 439
Bhili 9.7
Bhojpuri 9.7
Bhumij 29.2
Biafada 3.2.1
Biak 31.3.2 p. 183
Biatah 31.6.1
Bichelamar, see Bislama 206
Bicolano (Central), Ethn., see Bikol 31.6.3
Bidasoa Valley Basque, see Basque (Bidasoa Valley) 8
Bidiya 7.2.2
Bidiyo, Ethn. for Bidiya 7.2.2
Big Nambas, see Nambas (Big) 31.3.1
Bikol 31.6.3
Bila, Ethn., see Bira 3.3.1
Bilaan 31.6.7
Bilen, Ethn. for Bilin 7.3.2
Bilin 7.3.2
Bilinara 84.17
Biloxi 104 p. 451
Bilua 82
Bima 31.2
Bimoba 3.5
Binandere 57.4
Binanderean 57.4
Binga 6.2.1
Bini 3.3.4
Bininj Gun-Wok 84.7.2 pp. 127, 130, 134, 135, 422
Binukid 31.6.8
Binumarien 57.8
Bira 3.3.1
Bira, see Bira 3.3.1
Biri 84.17
Birobidjan Russian-Chinese Pidgin, see Russian-Chinese Pidgin (Birobidjan) 206
Birom 3.3.9 p. 295
Birri 6.2.1
Bisa 3.10.1 p. 58
Bischari, see Beja 7.3.1
Bislama 206
Bissa, Ethn. for Bisa 3.10.1
Bisu 26.2.4
Biu-Mandara 7.2.1
Blaan (Koronadal), Ethn., see Bilaan 31.6.7
Blaan (Sarangani), Ethn., see Bilaan 31.6.7
Blackfoot 89.1 p. 342
Blowo Dan, see Dan (Blowo) 3.10.1
Boazi 59.1
Bobangi 3.3.1
Bobo Fing 3.10.2
Bobo Madaré (Northern), Ethn., see Bobo Fing 3.10.2
Bobo Madaré (Southern), Ethn., see Bobo Fing 3.10.2
Bodega Miwok, see Miwok (Bodega) 113.7.2
Bodic 26.2.3 p. 95
Bodo 26.2.2
Boiken 52.1
Boikin, Ethn. for Boiken 52.1
Bokar 26.2.9
Boko 3.10.1 p. 226
Bole 7.2.4
Bolia 3.3.1
Bolivian Quechua, see Quechua (Bolivian) 169
Bologna Italian, see Italian (Bologna) 9.9
Bonan, Ethn. for Baonan 11.1
Bondo, Ethn. for Remo 29.2
Bongo 6.2.1
Bongo-Bagirmi 6.2.1
Bongu 57.14
Bontok 31.6.4
Bontok (Central), Ethn. for Bontok 31.6.4
Bora 148.2
Boraana Oromo, see Oromo (Boraana) 7.3.3
Boraana, see Oromo (Boraana) 7.3.3
Boran 148.1
Borana-Arsi-Guji, Ethn., see Oromo (Boraana) 7.3.3
Border 46
Bori 26.2.9
Borneo 31.6.1
Borno Nigerian Arabic, see Arabic (Borno Nigerian) 7.6
Boro (Omotic), Ethn. for Shinassha 7.5
Boro, see Bodo 26.2.2

Bororo 177.1 pp. 103, 478
Borôro, Ethn. for Bororo 177.1
Boruca 135.8
Bosavi 66
Bosavi, see Kaluli 66
Bosnian 9.10
Botin 53.2
Bozo (Tiéyaxo), Ethn. for Bozo (Tigemaxo) 3.10.2
Bozo (Tigemaxo) 3.10.2
Brabantic Dutch, see Dutch (Brabantic) 9.5
Brahui 25.2 pp. 87, 143, 151, 178, 230
Brao 29.1.2
Brazilian Sign Language, Ethn. for Língua de Sinais Brasileira 207
Breton 9.4 p. 475
Bribri 135.8
British Sign Language 207 pp. 558, 559, 561, 565
Broken 206
Brokskat 9.7
Bru (Eastern) 29.1.3
Bru (Western) 29.1.3
Bube, Ethn. for Bubi 3.3.1
Bubi 3.3.1
Budukh 20.1.3
Buduma 7.2.1
Buginese, see Bugis 31.6.9
Bugis 31.6.9
Buglere 135.3
Bugun 26.2.9
Bugurdz, see Romani (Bugurdzi) 9.7
Bugurdzi Romani, see Romani (Bugurdzi) 9.7
Buin 81
Bujeba 3.3.1
Bukar Sadong, Ethn., see Mentuh Tapuh 31.6.1
Bukidnon, see Manobo (Western Bukidnon) 31.6.8
Bukiyip, Ethn., see Arapesh 51.1
Bukiyip, see Arapesh 51.1
Bukusu 3.3.1
Bulaka River 62
Bularnu 84.17
Bulgarian 9.10 pp. 151, 315, 334
Buli (in Ghana) 3.5
Buli (in Indonesia) 31.3.2
Buli, Ethn. for Buli (in Ghana) 3.5
Buli, Ethn. for Buli (in Indonesia) 31.3.2
Bulu 3.3.1
Buma 31.3.1
Bunaba, Ethn. for Bunuba 84.1
Bundjalung, see Bandjalang 84.17
Bunu (Younuo), Ethn., see Punu 27
Bunuba 84.1 p. 526
Bunuban 84.1
Bura-Mabang, see Maba 6.8
Bura-Pabir 7.2.1
Burarra 84.2
Burarran 84.2
Burgenland Romani, see Romani (Burgenland) 9.7
Buriat 11.1 pp. 42, 523
Buriat (Mongolia), Ethn., see Buriat 11.1
Burji 7.3.3
Burkina Faso Fula, see Fula (Burkina Faso) 3.2.1
Burmese 26.2.4 pp. 83, 94, 147, 186, 187, 250, 271, 399, 451, 458, 459, 463
Burmese-Lolo 26.2.4
Buru 31.2
Burum 57.10
Burum-Mindik, Ethn. for Burum 57.10
Burunge 7.3.4 pp. 183, 471
Burushaski 22 pp. 42, 43, 67, 118, 127, 199, 222, 235, 242, 269, 327a
Busa 3.10.1
Bushonga, see Bushoong 3.3.1
Bushoong 3.3.1
Bwe Karen, see Karen (Bwe) 26.2.6
Byangsi, Ethn. for Byansi 26.2.3
Byansi 26.2.3
Byelorussian, see Belorussian 9.10
Cabécar 135.8
Cacua 146
Cacua-Nukak 146
Caddo 103
Caddoan 103 p. 78
Cahita 108.2
Cahuapanan 165
Cahuilla 108.6 pp. 43, 62, 438, 443, 522, 523, 526
Cajamarca Quechua, see Quechua (Cajamarca) 169
Cakchiquel 132 p. 438
Cakchiquel (Southern), Ethn., see Cakchiquel 132
Cambodian, see Khmer 29.1.5
Cameroonian Fula, see Fula (Cameroonian) 3.2.1
Camling 26.2.3
Campa (Axininca) 174 pp. 178, 523, 548
Campa Pajonal Asheninca 174 p. 523
Camsá 141
Canala, see Xârâcùù 31.3.1
Canamarí 185
Canary Islands Spanish, see Spanish (Canary Islands) 9.9
Candoshi 163
Candoshi-Shapra, Ethn. for Candoshi 163
Canela 177.2
Canela-Krahô 177.2 pp. 22, 42, 43, 162, 163, 287, 399, 400, 438, 510, 514,
Canela, Ethn., see Canela-Krahô 177.2
Cantonese 26.1 pp. 83, 127, 227, 246, 430, 522, 554, 565
Capanahua 168 p. 66
Car Nicobarese, see Nicobarese (Car) 29.1.7
Carapana 145 p. 523
Carib 175 p. 123
Carib (De'kwana) 175
Carib (Island), see Garífuna 174
Cariban 175 pp. 122, 323
Carijona 175 p. 526
Carijona, Ethn., see Hianacoto 175
Carijona, Ethn., see Umaua 175
Cariri, see Kariri 177.5
Carolinian 31.3.1
Carrier 87.1 p. 523
Cashibo 168
Cashibo-Cacataibo, Ethn., see Cashibo 168
Cashinahua 168 p. 548
Casiguran Dumagat, see Dumagat (Casiguran) 31.6.4
Casino Bandjalang, see Bandjalang (Casino) 84.17
Catalan 9.9
Catalan-Valencian-Balear, Ethn. for Catalan 9.9
Catawba 104
Cathlamet, see Chinook (Upper) 113.1
Catio 136
Cavineña 186
Cayapa 140
Cayubaba, Ethn. for Cayuvava 188
Cayuga 91.1
Cayuvava 188 pp. 139, 179, 211, 223, 243
Cebuano 31.6.3 pp. 330, 435
Celtic 9.4 pp. 58, 199, 270, 291, 331, 335, 359
Cèmuhî 31.3.1
Cemuhî, Ethn. for Cèmuhî 31.3.1
Central Agta, see Agta (Central) 31.6.4
Central Cushitic 7.3.2
Central Dravidian 25.1
Central Khoisan 1.1
Central Kurdish, see Kurdish (Central) 9.8
Central Malayo-Polynesian 31.2
Central Nahuatl, see Nahuatl (Central) 108.1
Central Pomo, see Pomo (Central) 117.3
Central Salish 110.2
Central Sierra Miwok, see Miwok (Central Sierra) 113.7.2
Central Sudanic 6.2 pp. 86, 339
Central Tarahumara, see Tarahumara (Central) 108.7
Central Yup'ik, see Yup'ik (Central) 86
Central-South Saami, see Saami (Central-South) 10.1.1
Cerma, Ethn. for Kirma 3.5
Ch'ol (Tila), Ethn., see Chol 132
Ch'ol (Tumbalá), Ethn., see Chol 132
Chachi, Ethn. for Cayapa 140
Chácobo 168
Chacobo, Ethn. for Chácobo 168
Chadic 7.2 pp. 18, 35, 83, 182, 283
Chaga 3.3.1
Chagga, Ethn. for Chaga 3.3.1
Chaha 7.6
Chai 6.3.8 pp. 358, 522
Chaladsch, see Khalaj 11.3
Chalcatongo Mixtec, see Mixtec (Chalcatongo) 124.4
Chaldean (Modern) 7.6
Chaldean (Neo-Aramaic), Ethn., see Neo-Aramaic (Amadiya) 7.6
Chaldean Neo-Aramaic, Ethn. for Chaldean (Modern) 7.6
Cham (Eastern) 31.6.10
Cham (Western) 31.6.10
Chamalal 20.1.1
Chambri 53.4
Chamorro 31.6.2 pp. 47, 86, 94, 102, 103, 163, 255, 263, 401, 426, 427, 438, 467
Chang 26.2.2
Changzhou Wu, see Wu (Changzhou) 26.1
Chantyal 26.2.3 p. 226
Chaouia Berber, see Berber (Chaouia) 7.1
Chaouia, Ethn. for Berber (Chaouia) 7.1
Chaozhou 26.1
Chapacura-Wanhan 182
Chasta Costa 87.1
Chasu, see Pare 3.3.1
Chatino (Nopala) 124.8
Chatino (Sierra Occidental) 124.8
Chatino (Sierra Occidental), Ethn., see Chatino (Yaitepec) 124.8
Chatino (Tataltepec) 124.8
Chatino (Yaitepec) 124.8
Chaudangsi 26.2.3
Chavacano 206
Chayahuita 165 p. 526
Chayuco Mixtec, see Mixtec (Chayuco) 124.4
Chechen 20.2 pp. 102, 106, 299, 438
Chehalis (Upper) 110.5
Cheke Holo 31.3.1
Chemakum 111
Chemehuevi 108.5 pp. 410, 522, 523
Chepang 26.2.3 p. 418
Cheremis (Eastern), see Mari (Meadow) 10.1.1
Cherokee 91.2 pp. 15, 155, 569
Chevak Yup'ik, see Yup'ik (Chevak) 86
Cheyenne 89.1
Chibcha, Ethn. for Muisca 135.2
Chibchan 135
Chibchan Motilón, see Motilón (Chibchan) 135.5
Chibchan Proper 135.2
Chicahuaxtla Trique, see Trique (Chicahuaxtla) 124.4
Chichewa 3.3.1 p. 127
Chichimec 124.2
Chichimeca-Jonaz 124.2 pp. 238, 239, 243
Chickasaw 94 p. 22
Chicomuceltec 132
Chiga, Ethn. for Nkore-Kiga 3.3.1
Chimakuan 111
Chimalapa Zoque, see Zoque (Chimalapa) 130
Chimariko 117.1 p. 548
Chimbu 57.5
Chimila 137
Chimúan 166
Chin (Asho), Ethn. for Hyow 26.2.7
Chin (Bawm), Ethn. for Bawm 26.2.7
Chin (Haka), Ethn. for Lai 26.2.7
Chin (Khumi), Ethn. for Khumi 26.2.7
Chin (Mara) 26.2.7
Chin (Paite), Ethn. for Paite 26.2.7
Chin (Siyin) 26.2.7
Chin (Tedim), Ethn. for Chin (Tiddim) 26.2.7
Chin (Thado), Ethn. for Thadou 26.2.7
Chin (Tiddim) 26.2.7
Chinantec (Comaltepec) 124.3
Chinantec (Lealao) 124.3 pp. 43, 127, 131, 175, 195, 299, 354
Chinantec (Palantla) 124.3
Chinantec (Quiotepec) 124.3 pp. 50, 370
Chinantec (San Felipe Usila) 124.3
Chinantec (Sochiapan) 124.3 p. 43
Chinantec (Tepetotutla) 124.3
Chinantecan 124.3 p. 283
Chinanteco (Comaltepec), Ethn. for Chinantec (Comaltepec) 124.3
Chinanteco (Lealao), Ethn. for Chinantec (Lealao) 124.3
Chinanteco (Palantla), Ethn. for Chinantec (Palantla) 124.3
Chinanteco (Quiotepec), Ethn. for Chinantec (Quiotepec) 124.3

Chinanteco (Sochiapan), Ethn. for Chinantec (Sochiapan) 124.3
Chinanteco (Usila), Ethn. for Chinantec (San Felipe Usila) 124.3
Chinese 26.1 pp. 51, 162, 164, 390, 394, 569
Chinese (Hakka), Ethn. for Hakka 26.1
Chinese (Mandarin), Ethn. for Mandarin 26.1
Chinese (Mandarin), Ethn., see Kunming 26.1
Chinese (Min Dong), Ethn., see Fuzhou 26.1
Chinese (Min Nan), Ethn., see Chaozhou 26.1 p. 554
Chinese (Min Nan), Ethn., see Taiwanese 26.1 p. 554
Chinese (Min Nan), Ethn., see Xiamen 26.1 p. 554
Chinese (Wu), Ethn., see Wu (Changzhou) 26.1
Chinese (Yue), Ethn. for Cantonese 26.1
Chinese Sign Language 207 pp. 558, 559, 561
Chinese, see Cantonese 26.1
Chinese, see Chaozhou 26.1
Chinese, see Fuzhou 26.1
Chinese, see Hakka 26.1
Chinese, see Hokkien 26.1
Chinese, see Kunming 26.1
Chinese, see Mandarin 26.1
Chinese, see Taiwanese 26.1
Chinese, see Wu (Changzhou) 26.1
Chinese, see Xiamen 26.1
Chingoni, see Ngoni 3.3.1
Chingpaw, see Jingpho 26.2.5
Chinook (Lower) 113.1
Chinook (Upper) 113.1 p. 548
Chinook, Ethn. for Chinook (Lower) 113.1
Chinookan 113.1 pp. 548, 549
Chipewyan 87.1 pp. 78, 83, 523
Chippewa (Red Lake and Pillager) 89.1
Chippewa, Ethn., see Chippewa (Red Lake and Pillager) 89.1
Chippewa, Ethn., see Ojibwe (Minnesota) 89.1
Chiquihuitlán Mazatec, see Mazatec (Chiquihuitlán) 124.7
Chiquitano, Ethn. for Chiquito 195
Chiquito 195
Chiricahua Apache, see Apache (Chiricahua) 87.1
Chiriguano 176.6 p. 523
Chitimacha 98
Chittagong Bengali, see Bengali (Chittagong) 9.7
Chittagonian, Ethn. for Bengali (Chittagong) 9.7
Chocho 124.7
Chochoteco, Ethn. for Chocho 124.7
Choco 136
Choctaw 94 pp. 114, 438, 526
Chol 132
Cholon 167
Chon 204
Chon Proper 204.1
Chontal (Highland) 129 p. 548
Chontal (Huamelultec Oaxaca) 129 p. 573
Chontal (Tabasco), Ethn. for Chontal Maya 132
Chontal de Oaxaca (Costa), Ethn. for Chontal (Huamelultec Oaxaca) 129
Chontal de Oaxaca (Sierra), Ethn. for Chontal (Highland) 129
Chontal Maya 132
Chorote 198
Chorote (Iyo'wujwa), Ethn., see Chorote 198
Chorote (Iyojwa'ja), Ethn., see Chorote 198
Chortí 132
Chrau 29.1.2
Chru 31.6.10
Chuave 57.5
Chuj 132
Chuj (San Mateo Ixtatán), Ethn., see Chuj 132
Chuj (San Sebastián Coatán), Ethn., see Chuj 132
Chukchi 14.1 pp. 43, 91, 94, 115, 126, 210, 223, 230, 272, 303, 351, 438, 439, 439a, 530, 548
Chukot, Ethn. for Chukchi 14.1
Chukotko-Kamchatkan 14 p. 548
Chulupí 198 p. 34
Chulym 11.3
Chumash 123
Chumash (Barbareño) 123 pp. 198, 422
Chumash (Ineseño) 123
Chumburung 3.9
Chuukese 31.3.1
Chuvash 11.3 pp. 70, 71, 146, 548
Cibemba, see Bemba 3.3.1
CiLuba 3.3.1
Cimambwe, see Mambwe 3.3.1
Cinsenga, see Nsenga 3.3.1
Circassian (East), see Kabardian 19
Circassian (West), see Adyghe (Abzakh) 19
Circassian (West), see Adyghe (Shapsugh) 19
Circassian (West), see Adyghe (Temirgoy) 19
Citak (Tamnim), Ethn., see Asmat 57.2
Clallam 110.2 p. 43
Coahuiltecan 101
Coahuilteco 101
Coast Tsimshian, see Tsimshian (Coast) 113.6
Coatlán Mixe, see Mixe (Coatlán) 130
Coatzospan Mixtec, see Mixtec (Coatzospan) 124.4
Cocama 176.6
Cocama-Cocamilla, Ethn. for Cocama 176.6
Cocama-Cocamilla, Ethn., see Tupi 176.6
Cochabamba Quechua, see Quechua (Cochabamba) 169
Cocopa 117.7 pp. 523, 526, 554
Coeur d'Alene 110.3 p. 422
Cofán 155
Cogui 135.1
Colac 84.17
Colorado, Ethn. for Tsafiki 140
Columbia-Wenatchi 110.3
Comaltepec Chinantec, see Chinantec (Comaltepec) 124.3
Comanche 108.5 pp. 222, 438, 522, 523
Comecrudan 102
Comecrudo 102 pp. 523, 526
Comorian 3.3.1
Comox 110.2
Coos (Hanis) 114.2 pp. 346, 382
Coos, Ethn. for Coos (Hanis) 114.2
Coosan 114.2
Copainalá Zoque, see Zoque (Copainalá) 130
Copala Trique, see Trique (Copala) 124.4
Coptic 7.4 p. 362
Cora 108.3 pp. 415, 522
Corachol 108.3
Cornish 9.4
Costanoan 113.7.1
Costanoan (Northern), Ethn., see Costanoan 113.7.1
Costanoan (Southern), Ethn., see Costanoan 113.7.1
Costanoan (Southern), Ethn., see Mutsun 113.7.1
Cou, see Tsou 31.5
Cree (Eastern) 89.1
Cree (Northern East), Ethn., see Cree (Eastern) 89.1
Cree (Plains) 89.1 pp. 87, 90, 94, 127, 131, 151, 203, 406, 435, 438, 569
Cree (Southern East), Ethn., see Cree (Eastern) 89.1
Creek 94
Creoles and Pidgins 206
Crimean Tatar 11.3
Crimean Turkish, Ethn. for Crimean Tatar 11.3
Crioulo (Upper Guinea), Ethn. for Guinea Bissau Crioulo 206
Croatian, see Serbian-Croatian 9.10
Cross River 3.3.2
Crow 104
Cua 29.1.2
Cubeo 145 pp. 79, 83, 523
Cuiba 142
Cuica 150
Cuicatec 124.4
Cuicateco (Tepeuxila), Ethn., see Cuicatec 124.4
Cuicateco (Teutila), Ethn., see Cuicatec 124.4
Cuitlatec 127
Cuiva, see Cuiba 142
Culina 184
Cuna, see Kuna 135.4
Cupeño 108.6
Curripaco 174
Cushitic 7.3 pp. 83, 203, 387
Cuzco Quechua, see Quechua (Cuzco) 169
Cypriot Greek, see Greek (Cypriot) 9.6
Czech 9.10 pp. 63, 522, 523, 573
Da'a 31.6.9
Daasanach, Ethn. for Dhaasanac 7.3.3
Daba 7.2.1
Dabida 3.3.1
Dadibi 71.2
Dafla, see Nishi 26.2.9
Daga 74 pp. 94, 123, 302, 415, 474, 475
Dagaare 3.5 p. 183
Dagaare (Southern), Ethn., see Dagaare 3.5
Dagan 74
Dagbani 3.5 pp. 83, 262, 427, 503
Daghestanian 20.1
Dagur 11.1 p. 50
Dahalo 7.3.4 pp. 35, 82, 83, 98
Daju 6.3.1
Daju (Dar Fur) 6.3.1 p. 58
Daju (Dar Sila), Ethn. for Sila 6.3.1
Dakota 104
Dakota, see Lakhota 104
Daly 84.3
Damana 135.1
Damin p. 573
Dan 3.10.1
Dan (Blowo) 3.10.1
Dan, Ethn., see Dan (Blowo) 3.10.1
Dangaléat (Western) 7.2.2
Dangaléat, Ethn., see Dangaléat (Western) 7.2.2
Dangme, see Gã 3.9
Dani 57.6
Dani (Lower Grand Valley) 57.6 pp. 42, 123, 178, 342, 399, 415
Danish 9.5 pp. 155, 311
Dar Fur Daju, see Daju (Dar Fur) 6.3.1
Darai 9.7 p. 54
Dargi, see Dargwa 20.1.2
Dargwa 20.1.2
Dari 9.8
Darling, Ethn. for Paakantyi 84.17
Darmiya 26.2.3
Dätiwuy Dhuwal, see Dhuwal (Dätiwuy) 84.17
Datooga 6.3.5
Daur, Ethn. for Dagur 11.1
Dâw 173
Day 3.1
Dayak (Ngaju), see Oloh Mangtangai 31.6.1
Dayak (Ngaju), see Pulopetak 31.6.1
Dazaga, Ethn., see Tubu 6.9
De'kwana Carib, see Carib (De'kwana) 175
Dedua p. 83
Defaka 3.6 p. 126
Defoid 3.3.3
Degema 3.3.4
Degexit'an 87.1 p. 523
Dehu, Ethn. for Drehu 31.3.1
Delaware, see Munsee 89.1
Delaware, see Unami 89.1
Dení 184
Deori, Ethn. for Deuri 26.2.2
Dera 47 p. 38
Dera, Ethn. for Kanakuru 7.2.4
Desano 145 p. 523
Deti 1.1 pp. 35, 82
Deuri 26.2.2
Deutsche Gebärdensprache 207 pp. 558, 559, 560, 561
Dhaasanac 7.3.3
Dhangu, Ethn., see Gaalpu 84.17
Dharawal 84.17
Dhargari 84.17
Dharumbal 84.17
Dhegiha, see Omaha 104
Dhimal 26.2.3
Dhivehi 9.7
Dholuo, see Luo 6.3.5
Dhurga 84.17
Dhuwal (Dätiwuy) 84.17
Dhuwal, Ethn., see Dhuwal (Dätiwuy) 84.17
Dhuwal, Ethn., see Djapu 84.17
Dida (Lakota), Ethn., see Vata 3.8
Didinga 6.3.8
Dido, Ethn. for Tsez 20.1.1
Diegueño (Mesa Grande) 117.7 pp. 23, 242, 282, 355, 366, 523, 526, 548
Dieri, Ethn. for Diyari 84.17
Digaro 26.2.9
Dii, Ethn. for Yag Dii 3.1
Dimasa 26.2.2
Dime 7.5
Dimili, see Zazaki 9.8
Dimli, Ethn. for Zazaki 9.8
Ding 3.3.1

Dinka 6.3.5 pp. 78, 110, 138, 210, 331, 339, 343, 411, 471
Dinka (Northeastern), Ethn., see Dinka 6.3.5
Diola-Fogny 3.2.1 pp. 58, 122, 530
Diyari 84.17 pp. 19, 319, 366, 438, 439, 447, 494
Dizi 7.5
Djabugay 84.17 p. 526
Djambarrpuyngu 84.17
Djamindjung, Ethn. for Jaminjung 84.9
Djapu 84.17
Djaru 84.17 pp. 438, 451
Djeebbana, Ethn. for Ndjébbana 84.15
Djeragan 84.4
Djinang 84.17
Djingili 84.20 pp. 62, 526
Djuka, see Ndyuka 206
Dobel 31.2
Dogon 3.4 p. 146
Dogon (Toro So) 3.4
Dogon, Ethn., see Dogon (Toro So) 3.4
Dogri 9.7
Doko 3.3.1
Dolakha Newari, see Newari (Dolakha) 26.2.3
Dolgan 11.3 p. 42
Domaaki, Ethn. for Dumaki 9.7
Domari 9.7 p. 335
Donegal Irish, see Irish (Donegal) 9.4
Dong 28.2
Dong (Northern), Ethn., see Dong 28.2
Dong (Northern), Ethn., see Kam (Zhanglu) 28.2
Dong (Southern), Ethn., see Dong 28.2
Dongo 3.1 p. 339
Dongolese Nubian, see Nubian (Dongolese) 6.3.6
Dongxiang, Ethn. for Santa 11.1
Doutai 38
Doyayo 3.1 p. 359
Dravidian 25 pp. 22, 43, 127, 130, 134, 135, 162, 163, 194, 267, 315, 415
Dravidian Parji, see Parji (Dravidian) 25.1
Drehu 31.3.1 pp. 83, 183, 399, 400
Drung, Ethn. for Dulong 26.2.11
Duala 3.3.1 p. 490
Duka 3.3.7
Dullay (Gollango) 7.3.3
Dulong 26.2.11
Duma 3.3.1
Dumagat (Casiguran) 31.6.4
Dumaki 9.7
Dumbea, Ethn. for Paita 31.3.1
Dumi 26.2.3 p. 318
Dumo 45.3 pp. 78, 183, 548
Duna 57.7
Duna-Bogaya 57.7
Duruwa, Ethn. for Parji (Dravidian) 25.1
Dutch 9.5 pp. 71, 158, 194, 195, 223, 287, 314, 318, 331, 434, 442, 467, 483, 491
Dutch (Berbice Creole), Ethn. for Berbice Dutch Creole 206
Dutch (Brabantic) 9.5
Dutch (Limburg) 9.5
Dutch (Zeeuws) 9.5
Dutch Sign Language, Ethn. for Nederlandse Gebarentaal 207
Dutch, Ethn., see Dutch (Brabantic) 9.5
Dutch, Ethn., see Dutch (Limburg) 9.5
Dyaabugay, Ethn. for Djabugay 84.17
Dyimini 3.5
Dyirbal 84.17 pp. 26, 438
Dyula 3.10.2 p. 572
Dzing, see Ding 3.3.1
East Bird's Head 33
East Bougainville 81
East Chadic 7.2.2
East Futuna, see Futuna (East) 31.3.1
East Geelvink Bay 37
East Greenlandic, see Greenlandic (East) 86
East Strickland 65
East Yugur, Ethn. for Shira Yughur 11.1
East-Central Xorasan Turkic, see Turkic (East-Central Xorasan) 11.3
Easter Island, see Rapanui 31.3.1
Eastern Aleut, see Aleut (Eastern) 86
Eastern Armenian, see Armenian (Eastern) 9.2
Eastern Bru, see Bru (Eastern) 29.1.3
Eastern Cham, see Cham (Eastern) 31.6.10
Eastern Cree, see Cree (Eastern) 89.1
Eastern Cushitic 7.3.3
Eastern Daly 84.3.2
Eastern Frisian, see Frisian (Eastern) 9.5
Eastern Highlands 57.8
Eastern Jebel 6.3.2
Eastern Kanjobal, see Kanjobal (Eastern) 132
Eastern Kayah Li, see Kayah Li (Eastern) 26.2.6
Eastern Libyan Arabic, see Arabic (Eastern Libyan) 7.6
Eastern Malayo-Polynesian 31.3
Eastern Mande 3.10.1
Eastern Mnong, see Mnong (Eastern) 29.1.2
Eastern Ojibwa, see Ojibwa (Eastern) 89.1
Eastern Pomo, see Pomo (Eastern) 117.3
Eastern Sko 45.1
Eastern Sudanic 6.3 pp. 331, 335
Ebira 3.3.8
Ecuadorean Quechua, see Quechua (Ecuadorean) 169
Edo, Ethn. for Bini 3.3.4
Edoid 3.3.4
Edolo 66
Efate (North), Ethn., see Nguna 31.3.1
Efik 3.3.2 pp. 27, 38, 523
Ega 3.9
Egyptian (Ancient) pp. 331, 335, 569
Egyptian Arabic, see Arabic (Egyptian) 7.6
Egyptian-Coptic 7.4
Eipo 57.17 p. 43
Eipomek, Ethn. for Eipo 57.17
Ejagham 3.3.1 p. 50
Ekagi, see Ekari 57.20
Ekari 57.20 pp. 146, 530, 531
Ekoka !Xun, see !Xun (Ekoka) 1.2
Ekoti 3.3.1
Eleman 72
Elseng 43
Emai-Iuleha-Ora, Ethn., see Ora (in Nigeria) 3.3.4
Embaloh 31.6.10 p. 438
Embera 136
Emberá (Northern), Ethn. for Embera 136
Embera Chami 136
Emberá-Catío, Ethn. for Catio 136
Emberá-Chamí, Ethn. for Embera Chami 136
Emberá-Saija, Ethn. for Epena Pedee 136
Emiliano-Romagnolo, Ethn. for Italian (Bologna) 9.9
Emmi 84.3.6
Enets 10.2
Enga 57.9 p. 43
Engan 57.9
Engenni 3.3.4
Enggano 31.6.10
English 9.5 pp. 10, 14, 15, 22, 26, 34, 38, 51, 54, 62, 78, 79, 82, 83, 94, 94, 102, 126, 127, 138, 142, 146, 147, 150, 151, 154, 158, 159, 162, 164, 170, 174, 182, 183, 190, 191, 194, 195, 206, 207, 214, 222, 223, 226, 227, 230, 231, 234, 238, 242, 246, 247, 250, 251, 254, 255, 258, 262, 266, 267, 269, 270, 271, 272, 282, 287, 290, 294, 298, 302, 303, 306, 310, 315, 322, 330, 334, 342, 347, 350, 351, 354, 358, 362, 363, 366, 367, 370, 375, 378, 379, 382, 383, 386, 390, 394, 399, 400, 407, 410, 411, 419, 426, 427, 430, 431, 434, 435, 439, 442, 454, 463, 466, 470, 474, 478, 482, 486, 487, 490, 491, 494, 495, 502, 514, 518, 522, 526, 530, 531, 546, 568, 572
English (Old) p. 195
Enya 3.3.1
Epena Pedee 136 pp. 46, 66, 175, 178, 251, 306, 382, 383
Erromangan 31.3.1
Erzya Mordvin, see Mordvin (Erzya) 10.1.1
Erzya, Ethn. for Mordvin (Erzya) 10.1.1
Ese Ejja 186
Eskimo-Aleut 86 pp. 11, 22, 31, 43, 115, 203, 401, 407, 491, 569
Esmeraldeño 156
Essejja, see Ese Ejja 186
Esselen 122
Estonian 10.1.1 pp. 214, 215, 218, 295, 407, 415, 490
Etruscan p. 331
Etsako 3.3.4
Eudeve 108.2 pp. 523, 526
Even 11.2 p. 42
Evenki 11.2 pp. 90, 143, 146, 151, 163, 195, 306, 311, 415, 523
Evenki, Ethn., see Solon 11.2
Ewe 3.9 pp. 83, 127, 287, 323
Ewe (Anglo) 3.9
Éwé, Ethn. for Ewe 3.9
Éwé, Ethn., see Ewe (Anglo) 3.9
Ewondo 3.3.1
Eyak 87.2 pp. 78, 79
Fante, see Akan 3.9
Faroese 9.5
Farsi (Eastern), Ethn. for Dari 9.8
Farsi (Western), Ethn. for Persian 9.8
Fasu 57.13 p. 314
Fe'fe' 3.3.1
Fehan, see Tetun 31.2
Fer, see Kara (in Central African Republic) 6.2.1
Figuig Berber, see Berber (Figuig) 7.1
Figuig, see Berber (Figuig) 7.1
Fijian 31.3.1 pp. 26, 42, 54, 83, 86, 87, 90, 94, 102, 239, 271, 334, 338, 346, 443, 454
Fijian (Wayan) 31.3.1
Fijian (Western), Ethn., see Fijian (Wayan) 31.3.1
Filipino, see Tagalog 31.6.3
Finisterre-Huon 57.10
Finnic 10.1.1
Finnish 10.1.1 pp. 15, 50, 91, 118, 147, 195, 214, 215, 223, 270, 271, 315, 319, 350, 411, 454, 458, 459, 462, 486, 506, 548, 568
Finnish Sign Language 207 pp. 558, 559, 564, 565
Finno-Ugric 10.1 p. 215
Fiorentino Italian, see Italian (Fiorentino) 9.9
Fiote 3.3.1
Fitzroy Crossing Kriol, see Kriol (Fitzroy Crossing) 206
Foe 57.13
Foi, Ethn. for Foe 57.13
Folopa, Ethn. for Podopa 71.2
Fon-Gbe, Ethn. for Fongbe 3.9
Fon, see Fongbe 3.9
Fongbe 3.9
Fordata 31.2
Fore 57.8 p. 358
Fox 89.1
Frafra, Ethn. for Gurenne 3.5
Francisco León Zoque, see Zoque (Francisco León) 130
French 9.9 pp. 14, 22, 31, 46, 50, 51, 63, 87, 127, 130, 170, 171, 174, 175, 186, 187, 190, 230, 270, 302, 315, 326, 338, 350, 374, 400, 435, 450, 455, 459, 467, 491, 531, 548, 573
French (Haitian Creole), Ethn. for Haitian Creole 206
French (Old) p. 130
French Guianese Creole French, Ethn. for Guianese French Creole 206
French Sign Language, Ethn. for Langue des Signes Française 207
Frisian 9.5
Frisian (Eastern) 9.5
Frisian (Eastern), Ethn., see Frisian 9.5
Frisian (North) 9.5
Frisian (Northern), Ethn. for Frisian (North) 9.5
Frisian (Western) 9.5
Fu-Yu Kirghiz, see Kirghiz (Fu-Yu) 11.3
Ful (Liptako) 3.2.1
Fula (Burkina Faso) 3.2.1
Fula (Cameroonian) 3.2.1
Fula (Guinean) 3.2.1
Fula (Mauritanian) 3.2.1 pp. 572, 573
Fula (Nigerian) 3.2.1 pp. 127, 548
Fula (Senegal) 3.2.1
Fulani (Gombe) 3.2.1
Fulfulde (Adamawa), Ethn. for Fula (Cameroonian) 3.2.1
Fulfulde (Adamawa), Ethn., see Fulani (Gombe) 3.2.1
Fulfulde (Maasina) 3.2.1 pp. 130, 131
Fulfulde (Nigerian), Ethn. for Fula (Nigerian) 3.2.1
Fulfulde (Northeastern Burkina Faso), Ethn. for Fula (Burkina Faso) 3.2.1
Fulfulde (Northeastern Burkina Faso), Ethn., see Ful (Liptako) 3.2.1
Fulniô 177.8
Fur 6.4 pp. 147, 206, 323

Furu, Ethn. for Bagiro 6.2.1
Futuna (East) 31.3.1
Futuna-Aniwa 31.3.1 pp. 155, 159
Fuuta Jalon, Ethn. for Fula (Guinean) 3.2.1
Fuzhou 26.1
Fwâi, Ethn. for Po-Ai 31.3.1
Fyam, Ethn. for Fyem 3.3.9
Fyem 3.3.9 pp. 470, 471
Gã 3.9 p. 222
Ga, Ethn. for Gã 3.9
Ga'anda 7.2.1
Gaagudju 84.5 p. 250
Gaalpu 84.17
Gaam, Ethn. for Ingessana 6.3.2
Gadaba 29.2 p. 302
Gadaba (Bodo), Ethn., see Gadaba 29.2
Gade 3.3.8
Gadsup 57.8 pp. 38, 79
Gaelic (Irish), Ethn. for Irish 9.4
Gaelic (Irish), Ethn., see Irish (Donegal) 9.4
Gaelic (Irish), Ethn., see Irish (Munster) 9.4
Gaelic (Scots) 9.4
Gagadu, Ethn. for Gaagudju 84.5
Gagauz 11.3 p. 255
Gahri 26.2.3
Gahuku 57.8
Galela 32.3
Galgadungic p. 164
Galician 9.9
Gallong 26.2.9
Gambian Mandinka, see Mandinka (Gambian) 3.10.2
Gamilaraay 84.17
Gamo 7.5
Gamo-Gofa-Dauro, Ethn., see Gamo 7.5
Gamo-Gofa-Dawro, Ethn., see Kullo 7.5
Gamu, see Kamu 84.3.2
Ganda, see Luganda 3.3.1
Ganggalida, Ethn. for Yukulta 84.18
Gapapaiwa 31.3.1
Gapun 54
Gapun, see Taiap 54
Garawa, Ethn. for Garrwa 84.6
Garífuna 174 p. 182
Garo 26.2.2
Garrwa 84.6 p. 74
Garrwan 84.6
Garus 57.14
Gavião 176.2 pp. 43, 250, 379,
Gavião do Jiparaná, Ethn. for Gavião 176.2
Gawwada, Ethn. for Dullay (Gollango) 7.3.3
Gayardilt, Ethn. for Kayardild 84.18
Gayo 31.6.10
Gbagyi, Ethn., see Gwari 3.3.8
Gbari, Ethn., see Gwari 3.3.8
Gbaya (Northwest), Ethn. for Gbaya Kara 3.1
Gbaya (Southwest), Ethn., see Mbodomo 3.1
Gbaya Kara 3.1
Gbaya-Bossangoa, Ethn. for Gbeya Bossangoa 3.1
Gbaya, Ethn. for Kresh 6.2.2
Gbeya Bossangoa 3.1
Ge-Kaingang 177.2
Gela 31.3.1 pp. 410, 411
Gelao 28.1
Genoa Italian, see Italian (Genoa) 9.9
Georgian 21 pp. 35, 119, 222, 286, 299, 314, 318, 319, 398, 399, 400, 450, 466, 546, 548, 572
German 9.5 pp. 14, 30, 31, 34, 51, 71, 91, 95, 122, 123, 127, 130, 147, 158, 170, 171, 186, 187, 191, 194, 195, 215, 223, 266, 331, 334, 338, 339, 431, 434, 435, 458, 466, 467, 470, 491, 494, 546, 548
German (Appenzell) 9.5 p. 223
German (Bavarian) 9.5 p. 467
German (Berlin) 9.5
German (Bern) 9.5
German (Hannover) 9.5
German (Mansfeldisch) 9.5
German (Ostschweiz) 9.5 p. 223
German (Ripuarian) 9.5
German (Standard), Ethn. for German 9.5
German (Standard), Ethn., see German (Berlin) 9.5
German (Standard), Ethn., see German (Hannover) 9.5
German (Standard), Ethn., see German (Mansfeldisch) 9.5
German (Standard), Ethn., see German (Thuringian) 9.5
German (Standard), Ethn., see German (Timisoara) 9.5
German (Thurgau) 9.5 p. 223
German (Thuringian) 9.5
German (Timisoara) 9.5 p. 223
German (Upper Austrian) 9.5
German (Viennese) 9.5
German (Westphalian) 9.5
German (Zurich) 9.5
German Sign Language, Ethn. for Deutsche Gebärdensprache 207
Germanic 9.5 pp. 58, 78, 190, 291, 315, 319, 467
Gernica Basque, see Basque (Gernica) 8
Ghat Tuareg, see Tuareg (Ghat) 7.1
Ghat, see Tuareg (Ghat) 7.1
Ghodoberi, Ethn. for Godoberi 20.1.1
Ghotuo 3.3.4
Ghulfan, Ethn., see Nubian (Hill) 6.3.6
Gidabal 84.17 p. 67
Gidar 7.2.1
Gikuyu, Ethn. for Kikuyu 3.3.1
Gilaki 9.8
Gilyak, Ethn. for Nivkh 15
Gilyak, Ethn., see Nivkh (South Sakhalin) 15
Gimira 7.5 pp. 366, 407, 470
Gitksan 113.6
Gitxsan, Ethn. for Gitksan 113.6
Giziga 7.2.1
Giziga (North), Ethn., see Giziga 7.2.1
Giziga (South), Ethn., see Giziga 7.2.1
Goajiro 174 pp. 218, 523
Godié 3.8 p. 430
Godoberi 20.1.1 pp. 126, 255, 438
Goemai 7.2.4
Gogodala 67
Gogodala-Suki 67
Goilalan 57.11
Gojri 9.7
Gokana 3.3.2
Gola 3.2.2
Golin 57.5 p. 67
Gollango Dullay, see Dullay (Gollango) 7.3.3
Gombe Fulani, see Fulani (Gombe) 3.2.1
Gondi 25.3
Gondi (Northern), Ethn., see Gondi 25.3
Gondi (Southern), Ethn., see Gondi 25.3
Gonja, see Nkonya 3.9
Gooniyandi 84.1 pp. 42, 86, 147, 163, 306, 318, 399, 438, 451, 466, 467, 496
Gorontalo 31.6.9
Gorowa 7.3.4 p. 74
Gourmanchéma, Ethn. for Gurma 3.5
Gourmanchéma, Ethn., see Gurma (Togo) 3.5
Gourmantche, see Gurma 3.5
Great Andamanese 30.1 p. 78
Grebo 3.8 pp. 43, 131, 147, 179, 295, 343, 454, 547, 548
Grebo (Barclayville), Ethn., see Grebo 3.8
Greek 9.6 p. 58
Greek (Ancient) pp. 27, 115
Greek (Cypriot) 9.6
Greek (Modern) 9.6 pp. 43, 83, 91, 91, 106, 115, 147, 190, 223, 310, 315, 447, 451, 466, 502, 503, 548
Greek (Mycenean) p. 569
Greek Sign Language 207 p. 558
Greek, Ethn. for Greek (Modern) 9.6
Greek, Ethn., see Greek (Cypriot) 9.6
Greenlandic (East) 86
Greenlandic (South) 86
Greenlandic (West) 86 pp. 31, 42, 110, 142, 202, 298, 303, 310, 311, 407, 430, 438, 439, 474, 506, 514, 518, 519
Guadeloupe Creole 206
Guahiban 142
Guahibo 142
Guahibo, Ethn., see Sikuani 142
Guaicuruan 199
Guajajara 176.6 pp. 422, 475, 523
Guajajára, Ethn. for Guajajara 176.6
Guambiano 140
Guana 197
Guanano 145 p. 523
Guaque 175
Guaraní 176.6 pp. 46, 47, 238, 239, 298, 523
Guaraní (Paraguayan), Ethn. for Guaraní 176.6
Guaraní (Western Argentine), Ethn. for Chiriguano 176.6
Guarayú, see Tapieté 176.6
Guarequena, Ethn. for Warekena 174
Guarijío 108.7
Guató 177.3
Guayabero 142
Guayakí, see Aché 176.6
Guaymi 135.3
Guaymi, see Ngäbere 135.3
Gude 7.2.1 pp. 358, 395, 435
Guduf 7.2.1
Guere 3.8
Gugada 84.17
Gugu-Yalanji, see Kuku-Yalanji 84.17
Guguyimidjir, Ethn. for Guugu Yimidhirr 84.17
Guhu-Samane 57.4
Guianese French Creole 206
Guinea Bissau Crioulo 206
Guinean Fula, see Fula (Guinean) 3.2.1
Gujarati 9.7
Gujari, Ethn. for Gojri 9.7
Gula (in Central African Republic) 6.2.1
Gula Iro 3.1
Gula, Ethn. for Gula (in Central African Republic) 6.2.1
Gulf Arabic, see Arabic (Gulf) 7.6
Gumatj 84.17
Gumawana 31.3.1 p. 435
Gumbaynggir 84.17 p. 478
Gumuz 6.5
Gunbalang 84.7.2
Gunin 84.21
Gunu 3.3.1
Gününa Küne 204.2
Gunwinggu, Ethn. for Bininj Gun-Wok 84.7.2
Gunwinygic 84.7.2
Gunwinyguan 84.7
Gunya 84.17
Gur 3.5 p. 343
Gurage (Soddo), Ethn. for Soddo 7.6
Gurage (West), Ethn., see Chaha 7.6
Gurage (West), Ethn., see Muher 7.6
Guragone, Ethn. for Gurr-goni 84.2
Gureng Gureng 84.17
Gurenne 3.5
Gurindji 84.17
Gurinji, Ethn. for Gurindji 84.17
Gurma 3.5 p. 522
Gurma (Togo) 3.5
Guro 3.10.1
Gurr-goni 84.2 pp. 142, 210
Gurung 26.2.3
Gurung (Eastern), Ethn., see Gurung 26.2.3
Gurung (Western), Ethn., see Gurung 26.2.3
Gusii 3.3.1
Guugu Yimidhirr 84.17
Gwari 3.3.8 p. 35
Gwich'in, see Kutchin 87.1
Gworok 3.3.9 p. 138
Hadza 2 p. 82
Haida 88 pp. 15, 226, 314
Haida (Northern) 88 pp. 63, 71
Haida (Northern), Ethn., see Haida 88
Haida (Southern), Ethn., see Haida 88
Haisla 109.1 p. 230
Haitian Creole 206 p. 459
Hakka 26.1
Halang 29.1.2
Halbi 9.7
Halia 31.3.1
Halkomelem 110.2 pp. 438, 439, 443, 443
Hamer 7.5 p. 35
Hamer-Banna, Ethn. for Hamer 7.5
Hamtai 57.1 p. 147
Hanga Hundi 52.1
Hani 26.2.4
Hanis Coos, see Coos (Hanis) 114.2
Hannover German, see German (Hannover) 9.5
Hano, Ethn. for Raga 31.3.1
Hanunóo 31.6.3
Hanunoo, Ethn. for Hanunóo 31.6.3

Harakmbet 171
Harar Oromo, see Oromo (Harar) 7.3.3
Harar, see Oromo (Harar) 7.3.3
Harari 7.6
Hare, see Slave 87.1
Harsusi 7.6
Haruai 56
Hassaniyya, Ethn. for Arabic (Bani-Hassan) 7.6
Hatam 32.1 pp. 142, 226, 250, 374, 390
Hausa 7.2.4 pp. 34, 35, 43, 127, 171, 182, 191, 262, 295, 455, 569
Havasupai 117.7
Havasupai-Walapai-Yavapai, Ethn., see Havasupai 117.7
Havasupai-Walapai-Yavapai, Ethn., see Hualapai 117.7
Havasupai-Walapai-Yavapai, Ethn., see Yavapai 117.7
Hawai'i Creole English, Ethn. for Hawaiian Creole 206
Hawaiian 31.3.1 pp. 54, 78, 138, 139, 178, 523, 526
Hawaiian Creole 206
Hawrami 9.8
Haya 3.3.1
Hayu 26.2.3
Hdi 7.2.1
Hebrew (Modern Ashenazic) 7.6
Hebrew (Modern) 7.6 pp. 86, 127, 170, 195, 226, 230, 250, 311, 430, 487, 568, 569, 572, 573
Hebrew, Ethn. for Hebrew (Modern) 7.6
Hebrew, Ethn., see Hebrew (Modern Ashenazic) 7.6
Hehe 3.3.1
Heiltsuk 109.1
Hemba 3.3.1
Herero 3.3.1
Hezhen, see Jurchen 11.2
Hianacoto 175
Hidatsa 104
Highland Chontal, see Chontal (Highland) 129
Hijazi Arabic, see Arabic (Hijazi) 7.6
Hiligaynon 31.6.3
Hill Mari, see Mari (Hill) 10.1.1
Hill Nubian, see Nubian (Hill) 6.3.6
Hindi 9.7 pp. 35, 122, 127, 147, 165, 250, 398, 399, 400, 431, 434, 496, 548
Hinukh 20.1.1
Hixkaryana 175 pp. 67, 131, 151, 239, 330, 375, 431, 462
Hixkaryána, Ethn. for Hixkaryana 175
Hmar 26.2.7
Hmong (White), see Hmong Daw 27
Hmong Daw 27
Hmong Njua 27 pp. 42, 94, 162, 374, 451
Hmong-Mien 27 p. 43
Ho 29.2
Hoava 31.3.1
Hocák, Ethn. for Winnebago 104
Hokan 117
Hokkien 26.1
Holoholo 3.3.1
Hondarribia Basque, see Basque (Hondarribia) 8
Hong Kong Sign Language 207 pp. 558, 559, 560, 561, 564
Hopi 108.4 pp. 38, 50, 70, 522
Houailou, see Ajië 31.3.1
Hre 29.1.2
Hua 57.8 p. 415
Hualapai 117.7
Huallaga Quechua, see Quechua (Huallaga) 169
Huambisa 162
Huamelultec Oaxaca Chontal, see Chontal (Huamelultec Oaxaca) 129
Huao, see Waorani 157
Huari, see Aikaná 174
Huarijío, Ethn. for Guarijío 108.7
Huarpe 201
Huastec 132 pp. 67, 83, 334
Huasteca Nahuatl, see Nahuatl (Huasteca) 108.1
Huasteco (San Francisco Chontla), Ethn., see Huastec 132
Huauchinango Nahuatl, see Nahuatl (Huauchinango) 108.1
Huautla Mazatec, see Mazatec (Huautla) 124.7
Huave (San Mateo del Mar) 131 pp. 22, 334
Huavean 131
Huba, Ethn. for Kilba 7.2.1
Huelel, see Esselen 122
Huichol 108.3 p. 522
Huitoto 148.2
Huitoto (Minica) 148.2
Huitoto (Minica), Ethn., see Huitoto 148.2
Huitoto (Muinane) 148.2
Huitoto (Murui) 148.2 p. 182
Huitoto (Nipode), Ethn. for Huitoto (Muinane) 148.2
Huitotoan 148
Hun-Saare, Ethn. for Duka 3.3.7
Hunde 3.3.1 pp. 111, 211, 374
Hungarian 10.1.2 pp. 15, 50, 147, 151, 187, 202, 203, 215, 226, 262, 287, 294, 295, 311, 323, 334, 398, 415, 486, 487, 491, 546, 548
Hunzib 20.1.1 pp. 83, 114, 170, 218, 219, 251, 382, 398, 399, 438, 470, 503
Hupa 87.1 pp. 523, 526
Hupda 173 pp. 222, 573
Hupdë, Ethn. for Hupda 173
Huron, see Wyandot 91.1
Hyow 26.2.7
I'saka 45.2
Iaai 31.3.1 pp. 50, 334
Iai, see Iaai 31.3.1
Iate, see Fulniô 177.8
Iatmul 52.1
Iau 38 p. 86
Iban 31.6.10
Ibanag 31.6.4
Ibibio 3.3.2
Ica, Ethn. for Ika 135.1
Icelandic 9.5 pp. 127, 202, 434
Icelandic (Old) p. 151
Icelandic Sign Language 207 pp. 558, 561
Ida'an, Ethn., see Begak-Ida'an 31.6.1
Idoma 3.3.5 p. 295
Idomoid 3.3.5
Idu 26.2.9
Iduna 31.3.1
Ifugao (Batad) 31.6.4 p. 395
Ifumu 3.3.1
Igbo 3.3.6 pp. 54, 123, 127, 195, 259, 400, 523
Igboid 3.3.6
Igede 3.3.5
Ignaciano 174
Igorot, see Bontok 31.6.4
Iha 34
Ijo (Kolokuma) 3.6 pp. 83, 463
Ijoid 3.6
Ik 6.3.3 p. 35
Ika 135.1 pp. 22, 123, 314, 323, 407, 415, 422
Ikalanga, see Kalanga 3.3.1
Ila 3.3.1
Illinois 89.1
Ilocano 31.6.4 pp. 114, 115, 371
Imbabura Quechua, see Quechua (Imbabura) 169
Imonda 46 pp. 94, 139, 203, 250, 262, 399, 400, 463
Inanwatan 59.2
Indian dialects Indo-Pakistani Sign Language, see Indo-Pakistani Sign Language (Indian dialects) 207
Indian Sign Language, Ethn. for Indo-Pakistani Sign Language (Indian dialects) 207
Indic 9.7 pp. 163, 194, 315
Indo-Aryan pp. 163, 315
Indo-European 9 pp. 35, 43, 58, 90, 91, 94, 95, 115, 119, 122, 127, 130, 135, 191, 194, 199, 207, 215, 243, 269, 269, 307, 335, 395, 415, 431, 487, 496, 502, 523, 548, 549
Indo-Pakistani Sign Language (Indian dialects) 207 pp. 558, 560, 561, 562
Indo-Pakistani Sign Language (Karachi dialect) 207 pp. 558, 560, 561, 562
Indonesian 31.6.10 pp. 38, 43, 127, 138, 218, 268, 283, 370, 522
Indonesian (Irianese) 31.6.10
Indonesian (Jakarta) 31.6.10
Indonesian, Ethn., see Indonesian (Irianese) 31.6.10
Indonesian, Ethn., see Indonesian (Jakarta) 31.6.10
Ineseño Chumash, see Chumash (Ineseño) 123
Ineseño, Ethn. for Chumash (Ineseño) 123
Inga 169 pp. 71, 214
Ingalik, see Degexit'an 86
Ingessana 6.3.2
Ingrian, Ethn. for Izhor 10.1.1
Ingush 20.2 pp. 35, 83, 119, 287, 299, 415
Interior Salish 110.3
International Sign 207 p. 558
Inukititut (Eastern Canadian), Ethn., see Inuktitut (Quebec-Labrador) 86
Inukititut (Eastern Canadian), Ethn., see Inuktitut (Salluit) 86
Inukititut (Western Canadian), Ethn., see Inuktitut (Rankin Inlet) 86
Inukititut (Western Canadian), Ethn., see Kangiryuarmiut 86
Inuktitut (Greenlandic), Ethn., see Greenlandic (East) 86
Inuktitut (Greenlandic), Ethn., see Greenlandic (South) 86
Inuktitut (Greenlandic), Ethn., see Greenlandic (West) 86
Inuktitut (Quebec-Labrador) 86 pp. 234, 569
Inuktitut (Rankin Inlet) 86
Inuktitut (Salluit) 86
Iñupiaq 86 p. 526
Inupiatun (North Alaskan), Ethn. for Iñupiaq 86
Iowa-Oto, Ethn., see Oto 104
Ipurina, see Apurinã 174
Iquito 158
Irangi, see Langi 3.3.1
Iranian 9.8 pp. 199, 315, 387, 398
Iranian Armenian, see Armenian (Iranian) 9.2
Iranian Azari, see Azari (Iranian) 11.3
Irántxe, Ethn. for Iranxe 174
Iranxe 174
Iraqi Arabic, see Arabic (Iraqi) 7.6
Iraqw 7.3.4 pp. 70, 83, 122, 123, 171, 183, 230, 250, 291, 339, 387
Irarutu 31.3.2
Irianese Indonesian, see Indonesian (Irianese) 31.6.10
Irish 9.4 pp. 118, 199, 295, 310, 323, 330, 342, 451, 478, 479
Irish (Donegal) 9.4
Irish (Munster) 9.4
Irish Sign Language 207 pp. 558, 559, 561
Iro, see Gula Iro 3.1
Iroquoian 91
Isangu, see Sangu 3.3.1
Isekiri 3.3.3
Ishkashmi 9.8
Isirawa 39
Island Carib, see Garífuna 174
Isleta, see Tiwa (Southern) 105
Isnag 31.6.4
Isoko 3.3.4
Israeli Sign Language 207 pp. 558, 559
Issyk-Kul Kalmyk, see Kalmyk (Issyk-Kul) 11.1
Isthmus Zapotec, see Zapotec (Isthmus) 124.8
Italian 9.9 pp. 190, 467, 471, 506, 514, 518, 573
Italian (Bologna) 9.9
Italian (Fiorentino) 9.9
Italian (Genoa) 9.9
Italian (Napolitanian) 9.9
Italian (Turinese) 9.9
Italian Sign Language, Ethn. for Lingua Italiana dei Segni 207
Italian, Ethn., see Italian (Fiorentino) 9.9
Itawis 31.6.4
Itawit, Ethn. for Itawis 31.6.4
Itelmen 14.2 pp. 35, 147, 548
Itonama 187
Itzá, Ethn. for Itzaj 132
Itzaj 132 pp. 143, 407, 502
Iu Mien, Ethn. for Mien 27
Ivatan 31.6.4
Ivatan (Southern) 31.6.4
Ivatan, Ethn., see Ivatan (Southern) 31.6.4
Iwaidja 84.8
Iwaidjan 84.8
Iwam 52.5
Ixcatec 124.7
Ixcateco, Ethn. for Ixcatec 124.7
Ixil 132
Ixil (Nebaj), Ethn., see Ixil 132
Ixil (San Juan Cotzal), Ethn., see Ixil 132
Ixtenco Otomí, see Otomí (Ixtenco) 124.5

Ixtlan Zapotec, see Zapotec (Ixtlan) 124.8
Izhor 10.1.1
Izi 3.3.6 p. 455
Izi-Ezaa-Ikwo-Mgbo, Ethn., see Izi 3.3.6
Izon, Ethn. for Ijo (Kolokuma) 3.6
Jabêm 31.3.1
Jabutí 181
Jacaltec, see Jakaltek 132
Jacalteco (Eastern), Ethn., see Jakaltek 132
Jacalteco (Western), Ethn., see Jakaltek 132
Jad 26.2.3
Jagan, see Yahgan 205
Jahai 29.1.1 p. 573
Jakaltek 132 pp. 307, 323, 347, 411, 438
Jakarta Indonesian, see Indonesian (Jakarta) 31.6.10
Jamaican Creole 206
Jamamadi 184
Jamiltepec Mixtec, see Mixtec (Jamiltepec) 124.4
Jaminjung 84.9 pp. 190, 230
Jaminjungan 84.9
Jamul Tiipay, see Tiipay (Jamul) 117.7
Jangshung, Ethn. for Zhang-Zhung 26.2.3
Japanese 17 pp. 31, 58, 94, 150, 171, 186, 194, 195, 259, 263, 330, 400, 411, 450, 496, 565, 568, 569, 573
Japanese Sign Language Nihon Shuwa, see Nihon Shuwa (Japanese Sign Language) 207
Japanese Sign Language, Ethn. for Nihon Shuwa (Japanese Sign Language) 207
Japreria 175
Japrería, Ethn. for Japreria 175
Jaqaru 170 pp. 179, 198, 230, 423
Jarawa (in Andamans) 30.2
Jarawa (in Nigeria) 3.3.1
Jarawa, Ethn. for Jarawa (in Andamans) 30.2
Jarawa, Ethn. for Jarawa (in Nigeria) 3.3.1
Jarawara 184
Jaru, Ethn. for Djaru 84.17
Jaruára, Ethn. for Jarawara 184
Javaé 177.4 p. 526
Javanese 31.6.10 p. 162
Jebero 165
Jeh 29.1.2
Jehai, Ethn. for Jahai 29.1.1
Jeli 3.10.2
Jemez 105
Jiarong 26.2.13 p. 283
Jibbali 7.6
Jicaltepec Mixtec, see Mixtec (Jicaltepec) 124.4
Jicaque, see Tol 133
Jicarilla Apache, see Apache (Jicarilla) 87.1
Jiliapan Pame, see Pame 124.6
Jinghpo 26.2.5
Jingphaw, see Jingpho 26.2.5
Jingpho 26.2.5
Jingulu, see Djingili 84.20
Jino 26.2.4
Jinuo (Buyuan), Ethn., see Jino 26.2.4
Jinuo (Youle), Ethn., see Jino 26.2.4
Jirajaran 149
Jivaro 162 pp. 106, 526
Jivaroan 162
Johari 9.7
Jola-Fogny, Ethn. for Diola-Fogny 3.2.1
Jomang 3.7
Ju/'hoan, Ethn. for Ju|'hoan 1.2
Ju|'hoan 1.2 pp. 82, 127, 131, 147, 159, 362, 503
Juang 29.2
Juárez Zapotec, see Zapotec (Juárez) 124.8
Juat 84.17
Jugli 26.2.2
Jukun 3.3.9
Jukun Takum, Ethn. for Jukun 3.3.9
Jula, Ethn. for Dyula 3.10.2
Júma 176.6
Jur Luwo 6.3.5
Jur Mödö 6.2.1
Jur Modo, Ethn. for Jur Mödö 6.2.1
Jurchen 11.2
Juruna 176.6
Jurúna, Ethn. for Juruna 176.6
K'ekchí 132
Kabardian 19 pp. 35, 38, 83, 438, 471
Kabatei 9.8
Kabiyé 3.5
Kabui 26.2.7
Kabyle 7.1 pp. 486, 487
Kachari 26.2.2
Kacipo-Balesi, Ethn., see Baale 6.3.8
Kadai 28.1
Kadazan 31.6.1 p. 522
Kadazan (Coastal), Ethn., see Kadazan 31.6.1
Kadiwéu 199
Kadugli 5
Kaficho, Ethn. for Kefa 7.5
Kágaba, see Cogui 135.1
Kagoma 3.3.9
Kagulu, Ethn. for Kaguru 3.3.1
Kaguru 3.3.1
Kaian 53.3
Káidi-Kanembú, see Kanembu 6.9
Kaili 31.6.9
Kaili (Da'a), Ethn., see Da'a 31.6.9
Kaili (Ledo), Ethn. for Kaili 31.6.9
Kaingang 177.2
Kaingáng, Ethn. for Kaingang 177.2
Kainji 3.3.7
Kairiru 31.3.1 pp. 342, 370
Kaiwá 176.6
Kakchiquel, see Cakchiquel 132
Kaki Ae 72
Kala Lagaw Ya 84.17 p. 78
Kalam 57.14 pp. 222, 573
Kalami 9.7
Kalanga 3.3.1 p. 523
Kalapuya 116
Kalapuyan 116
Kalasha 9.7
Kalderash Romani, see Romani (Kalderash) 9.7
Kalenjin, Ethn., see Nandi 6.3.5
Kaliai-Kove 31.3.1
Kalispel 110.3 p. 526
Kalispel-Pend d'Oreille, Ethn. for Kalispel 110.3
Kalkatungu 84.17
Kalkutung, Ethn. for Kalkatungu 84.17
Kalmyk 11.1 p. 523
Kalmyk (Issyk-Kul) 11.1
Kalmyk-Oirat, Ethn., see Kalmyk 11.1
Kalmyk-Oirat, Ethn., see Kalmyk (Issyk-Kul) 11.1
Kalmyk-Oirat, Ethn., see Oirat 11.1
Kaluli 66
Kam (Zhanglu) 28.2
Kam-Tai 28.2
Kamã, Ethn. for Dâw 173
Kamaiurá 176.6
Kamang, Ethn. for Woisika 57.19
Kamano-Kafe 57.8
Kamano, Ethn. for Kamano-Kafe 57.8
Kamas, Ethn. for Kamass 10.2
Kamass 10.2
Kamayurá, Ethn. for Kamaiurá 176.6
Kamba 3.3.1
Kambera 31.2 pp. 147, 171
Kamberataro, Ethn. for Dera 47
Kambot 53.2
Kamchadal, see Itelmen 14.2
Kame.ntzá, see Camsá 141
Kami 3.3.1
Kamilaroi, Ethn. for Gamilaraay 84.17
Kamilaroi, Ethn., see Yuwaalaraay 84.17
Kammu, see Khmu' 29.1.8
Kamoro 57.2
Kamu 84.3.2
Kana 3.3.2 pp. 359, 466
Kanakuru 7.2.4
Kanamarí, Ethn. for Canamarí 185
Kanembu 6.9
Kangiryuarmiut 86
Kanjobal (Eastern) 132
Kanjobal (Western) 132
Kankanaey, Ethn. for Kankanay 31.6.4
Kankanay 31.6.4
Kannada 25.4 pp. 126, 127, 134, 163, 318, 383, 434
Kanum (Bädi) 63
Kanum (Ngkâlmpw) 63
Kanuri 6.9 pp. 126, 230, 323, 475
Kanuri (Central), Ethn., see Kanuri 6.9
Kanyok 3.3.1
Kapampangan 31.6.4 pp. 366, 367, 438, 482
Kapau, see Hamtai 57.1
Kapingamarangi 31.3.1
Kapriman, Ethn. for Sare 52.3
Kara (in Central African Republic) 6.2.1
Kara (in Papua New Guinea) 31.3.1
Kara, Ethn. for Kara (in Central African Republic) 6.2.1
Kara, Ethn. for Kara (in Papua New Guinea) 31.3.1
Karaboro (Western), Ethn., see Tenyer 3.5
Karachay-Balkar 11.3 pp. 299, 495, 496, 523
Karachi dialect Indo-Pakistani Sign Language, see Indo-Pakistani Sign Language (Karachi dialect) 207
Karadjeri 84.17
Karagas, Ethn. for Tofa 11.3
Karagash Noghay, see Noghay (Karagash) 11.3
Karaim 11.3 p. 299
Karajá 177.4
Karajá, Ethn., see Javaé 177.4
Karakalpak 11.3
Karamojong, Ethn. for Karimojong 6.3.5
Karanga 3.3.1
Karankawa 100 p. 526
Karao 31.6.4
Karata 20.1.1
Karbi, see Mikir 26.2.7
Karelian 10.1.1
Karen 26.2.6
Karen (Bwe) 26.2.6 p. 382
Karen (Pwo Eastern), Ethn., see Karen (Pwo) 26.2.6
Karen (Pwo Northern), Ethn. for Phlong 26.2.6
Karen (Pwo Western), Ethn., see Karen (Pwo) 26.2.6
Karen (Pwo) 26.2.6
Karen (S'gaw), Ethn. for Karen (Sgaw) 26.2.6
Karen (Sgaw) 26.2.6 pp. 83, 94
Karimojong 6.3.5
Karipuna do Guapore 168
Karipuná, Ethn., see Karipuna do Guapore 168
Kariri 177.5 p. 238
Kariri-Xocó, Ethn. for Kariri 177.5
Karitiâna 176.1
Karkar-Yuri 48
Karó (Arára) 176.4 p. 363
Karo Batak, see Batak (Karo) 31.6.10
Karok 118 pp. 46, 94, 123, 462, 463, 526, 546
Kartvelian 21 pp. 35, 119, 299, 314, 315, 319, 548
Kasem 3.5
Kashaya 117.3
Kashmiri 9.7 pp. 450, 523
Kashubian 9.10
Kashubian, Ethn., see Slovincian 9.10
Kata Kolok 207 pp. 558, 561
Katcha 5
Katcha-Kadugli-Miri, Ethn., see Kadugli 5
Katcha-Kadugli-Miri, Ethn., see Katcha 5
Kâte 57.10 pp. 43, 83, 143
Kate, see Kati (in West Papua, Indonesia) 57.18
Kathmandu Newari, see Newari (Kathmandu) 26.2.3
Kati (in Afghanistan) 9.7
Kati (in West Papua, Indonesia) 57.18
Kati, Ethn. for Kati (in Afghanistan) 9.7
Katla 3.7
Kato 87.1
Katu 29.1.3 p. 183
Katu (Eastern), Ethn., see Katu 29.1.3
Katu (Western), Ethn., see Katu 29.1.3
Katuic 29.1.3
Katukinan 185
Kaugel 57.5
Kaulong 31.3.1
Kaure 44
Kaurna 84.17 p. 526
Kawaiisu 108.5 pp. 522, 523
Kawesqar, see Qawasqar 203
Kayagar 58
Kayah (Eastern), Ethn. for Kayah Li (Eastern) 26.2.6
Kayah Li (Eastern) 26.2.6 pp. 42, 250, 251, 450, 451
Kayan (Baram) 31.6.1
Kayapó 177.2
Kayardild 84.18 pp. 74, 90, 115, 139, 202, 223, 250, 506
Kayu Agung 31.6.10
Kazakh 11.3 p. 523
Ke'o 31.2
Kebar 32.2

Kedang 31.2
Kefa 7.5
Kei 31.2
Kekchí, Ethn. for K'ekchí 132
Kela (Apoze) 31.3.1
Kela, Ethn., see Kela (Apoze) 31.3.1
Kelabit 31.6.1
Kele 31.3.1
Kemant 7.3.2
Kemtuik 41
Kenga 6.2.1 p. 411
Kenuzi-Dongola, Ethn., see Nubian (Dongolese) 6.3.6
Kenuzi-Dongola, Ethn., see Nubian (Kunuz) 6.3.6
Kenyah (Uma' Lung) 31.6.1
Kenyan Sign Language 207 pp. 558, 559
Kenyang 3.3.1
Kera 7.2.2 pp. 35, 171
Kerek 14.1 pp. 42, 291
Keres (Eastern), Ethn., see Keresan (Santa Ana) 106
Keres (Western), Ethn. for Acoma 106
Keresan 106 p. 35
Keresan (Santa Ana) 106 p. 526
Kesukuma, see Sukuma 3.3.1
Ket 13 pp. 27, 42, 58, 111, 235, 283, 303, 327a, 407, 571
Ketapang 31.6.10
Kete 3.3.1
Ketengban 57.17
Kewa 57.9 pp. 95, 314
Kewa (East), Ethn., see Kewa 57.9
Kewa (West), Ethn., see Kewa 57.9
Khakas 11.3 p. 523
Khalaj 11.3
Khaling 26.2.3
Khalkha 11.1 pp. 67, 164, 314, 523, 548
Kham 26.2.3
Kham (Gamale), Ethn., see Kham 26.2.3
Khamnigan Mongol, see Mongol (Khamnigan) 11.1
Khana, Ethn. for Kana 3.3.2
Khanty 10.1.2 pp. 42, 202, 415
Kharia 29.2 p. 178
Khasi 29.1.4 pp. 127, 178
Khasian 29.1.4
Khinalug 20.1.3
Khinalugh, Ethn. for Khinalug 20.1.3
Khmer 29.1.5 pp. 127, 139, 190, 447
Khmer (Central), Ethn. for Khmer 29.1.5
Khmu, Ethn. for Khmu' 29.1.8
Khmu' 29.1.8 p. 451
Khoekhoe 1.1 pp. 43, 82, 86, 115, 323
Khoisan 1 pp. 11, 35, 82, 127, 131, 183, 259, 287, 291, 355, 415, 573
Khorasani Turkish, Ethn., see Turkic (East-Central Xorasan) 11.3
Khorasani Turkish, Ethn., see Turkic (West Xorasan) 11.3
Khowar 9.7
Khumi 26.2.7 pp. 572, 573
Khvarshi 20.1.1
Kickapoo 89.1
Kihunde, see Hunde 3.3.1
Kikongo, see Kongo 3.3.1
Kikuyu 3.3.1
Kilba 7.2.1
Kildin Saami, see Saami (Kildin) 10.1.1
Kilivila 31.3.1 pp. 418, 451
Kiliwa 117.7
Kiluba 3.3.1
Kiluba-Hemba, see Hemba 3.3.1
Kim Mun, Ethn. for Mun 27
Kim, Ethn., see Kosop 3.1
Kimaama, Ethn. for Kimaghama 60
Kimaghama 60
Kimatuumbi, see Matuumbi 3.3.1
Kinande, see Nande 3.3.1
Kinga 3.3.1
Kinnauri 26.2.3
Kinyambo, see Nyambo 3.3.1
Kinyamwezi, see Nyamwezi 3.3.1
Kinyarwanda 3.3.1 pp. 427, 442, 443, 443a, 446
Kiowa 105 pp. 39, 94, 239, 303, 382, 438, 470, 548
Kiowa-Tanoan 105 pp. 315, 548
Kipea, see Kariri 177.5
Kiranti p. 95
Kire 53.5
Kirghiz 11.3 p. 523
Kirghiz (Fu-Yu) 11.3
Kirghiz, Ethn., see Kirghiz (Fu-Yu) 11.3
Kiribati 31.3.1 pp. 78, 111, 199, 235
Kirikiri 38
Kirma 3.5
Kisar 31.2
Kisi 3.2.2 pp. 35, 86, 90, 127, 139, 218, 339, 370
Kisi (Southern) 3.2.2
Kisi (Southern), Ethn., see Kisi 3.2.2
Kisonge, see Songe 3.3.1
Kissi (Northern), Ethn., see Kisi 3.2.2
Kitalinga, see Talinga 3.3.1
Kitja 84.4
Kitsai 103
Kituba 206 p. 523
Kiwai 68
Kiwai (Northeast), Ethn., see Kiwai 68
Kiwai (Southern), Ethn., see Kiwai 68
Kiwaian 68
Klamath 113.2
Klamath-Modoc 113.2
Klamath-Modoc, Ethn. for Klamath 113.2
Klao 3.8 p. 79
Ko (Winye) 3.5
Koasati 94 pp. 94, 95, 171, 306, 314, 406, 407, 415, 451
Kobon 57.14 pp. 122, 123, 158, 190, 206, 218, 367, 375, 407, 414, 446, 451, 518, 530, 531
Kodagu, Ethn. for Kodava 25.4
Kodava 25.4
Koegu 6.3.8 p. 407
Kofyar p. 183
Köggaba, see Cogui 135.1
Koh (Lakka) 3.1
Koho, Ethn., see Sre 29.1.2
Kohumono 3.3.2
Koiali (Mountain) 57.12
Koiari 57.12
Koiari (Grass), Ethn. for Koiari 57.12
Koiarian 57.12
Koita 57.12
Koitabu, Ethn. for Koita 57.12
Kok Borok, Ethn. for Kokborok 26.2.2
Kokata, Ethn. for Gugada 84.17
Kokborok 26.2.2
Kokni 9.7
Kokota 31.3.1
Kola 31.2
Kolami 25.1 pp. 127, 434
Kolami (Northwestern), Ethn., see Kolami 25.1
Kolami (Southeastern), Ethn., see Kolami 25.1
Kolana 57.19
Kolokuma Ijo, see Ijo (Kolokuma) 3.6
Kolopom 60
Kölsch, Ethn. for German (Ripuarian) 9.5
Koluri 9.8
Kolyma Yukaghir, see Yukaghir (Kolyma) 12
Kom 3.3.1
Koman 6.6
Kombai 57.3 pp. 366, 379
Kombio-Arapesh 51.1
Komi-Permyak 10.1.1 p. 67
Komi-Zyrian 10.1.1 pp. 314, 319
Komi-Zyrian, Ethn., see Yazva 10.1.1
Komo 6.6 p. 35
Komodo, see Bima 31.2
Konda 25.3
Konda-Dora, Ethn., see Konda 25.3
Kongo 3.3.1 pp. 147, 523
Kongo, Ethn., see Fiote 3.3.1
Konjo 31.6.9 p. 406
Konjo (Coastal), Ethn. for Konjo 31.6.9
Konkani 9.7 p. 131
Konkomba 3.5
Konkow 113.3
Konni 3.5
Konua 80
Konyagi 3.2.1
Koorete, Ethn. for Koyra 7.5
Korafe 57.4
Korana 1.1 pp. 110, 182, 450
Koranko 3.10.2
Kordofanian 3.7 p. 283
Korean 18 pp. 34, 35, 38, 94, 147, 174, 223, 250, 303, 496, 531, 565, 569
Korean Sign Language, Ethn. for South Korean Sign Language 207
Koreguaje 145 p. 523
Korku 29.2 pp. 142, 190, 483
Kormakiti Arabic, see Arabic (Kormakiti) 7.6
Koromfe 3.5 pp. 46, 58, 131, 147, 171, 262, 295, 346
Koromfé, Ethn. for Koromfe 3.5
Korowai 57.3 pp. 158, 371
Korowai (North), Ethn., see Korowai 57.3
Korowai, Ethn., see Korowai 57.3
Koryak 14.1
Kosarek Yale, see Yale (Kosarek) 57.17
Kosop 3.1
Kosraean 31.3.1
Kota 25.4
Koti, Ethn. for Ekoti 3.3.1
Kotia Oriya, see Oriya (Kotia) 9.7
Kotoko 7.2.1 pp. 35, 83
Kott 13
Kove, Ethn., see Kaliai-Kove 31.3.1
Koya 25.3
Koyra 7.5
Koyra Chiini 6.11 pp. 46, 195, 250, 346, 379, 386, 387
Koyraboro Senni 6.11 pp. 147, 171, 195, 295, 343, 371, 374, 379, 438, 443, 447, 462
Koyukon 87.1 p. 523
Kpan 3.3.9
Kpelle 3.10.2 pp. 35, 83
Kposo 3.9
Krahô 177.2
Krahô, Ethn., see Canela-Krahô 177.2
Kresh 6.2.2
Kriol (Fitzroy Crossing) 206
Kriol (Ngukurr) 206
Kriol, Ethn., see Kriol (Ngukurr) 206
Krisa 45.2
Krisa, Ethn. for I'saka 45.2
Kristang, see Malacca Creole 206
Krongo 5 pp. 118, 210, 282, 327, 350, 383, 415, 426, 438, 451
Kru 3.8 p. 343
Kru, see Klao 3.8
Krymchak 11.3
Kryts, Ethn. for Kryz 20.1.3
Kryz 20.1.3
Ksingmul 29.1.8
Kuala Lumpur Malay, see Malay (Kuala Lumpur) 31.6.10
Kualan 31.6.10
Kuanua, Ethn. for Tolai 31.3.1
Kugu Nganhcara 84.17
Kui (in India) 25.3
Kui (in Indonesia) 57.19
Kui, Ethn. for Kui (in India) 25.3
Kui, Ethn. for Kui (in Indonesia) 57.19
Kuki-Chin-Naga 26.2.7 pp. 95, 573
Kukna, Ethn. for Kokni 9.7
Kukú 6.3.5
Kuku-thaypan, see Thaypan 84.17
Kuku-Uwanh, Ethn. for Kugu Nganhcara 84.17
Kuku-Yalanji 84.17
Kulamanen 31.6.8
Kuliak 6.3.3
Kullo 7.5
Kulung 26.2.3
Kumak, Ethn. for Nelemwa 31.3.1
Kuman 57.5
Kumauni 9.7
Kumauni, Ethn., see Johari 9.7
Kumbainggar, Ethn. for Gumbaynggir 84.17
Kumiái, Ethn., see Diegueño (Mesa Grande) 117.7
Kumiái, Ethn., see Tiipay (Jamul) 117.7
Kumyk 11.3 pp. 298, 299, 523
Kuna 135.4
Kuna (Border), Ethn. for Kuna 135.4
Kunama 6.7 pp. 122, 327a
Kunbarlang, Ethn. for Gunbalang 84.7.2
Kundu, see Bakundu 3.3.1
Kung-Ekoka, Ethn. for !Xun (Ekoka) 1.2
Kunimaipa 57.11
Kunjen 84.17
Kunming 26.1

Kunuz Nubian, see Nubian (Kunuz) 6.3.6
Kunza 194
Kunza, Ethn. for Atacameño 194
Kuo, Ethn. for Koh (Lakka) 3.1
Kuot 79 p. 335
Kupsabiny, Ethn. for Sebei 6.3.5
Kuranko, Ethn. for Koranko 3.10.2
Kurdish (Central) 9.8 pp. 199, 568
Kurdish (Kurmanji), see Kurmanji 9.8
Kurmanji 9.8
Kurukh 25.2
Kurux, Ethn. for Kurukh 25.2
Kuskokwim (Upper), Ethn. for Upper Kuskokwim 87.1
Kusunda 24
Kutai 31.6.10
Kutchin 87.1 p. 523
Kutenai 112 pp. 39, 90, 139, 147, 159, 230, 283, 347, 351, 363, 391
Kutubuan 57.13
Kuuku Ya'u 84.17
Kuuku-Ya'u, Ethn. for Kuuku Ya'u 84.17
Kuvi 25.3
Kuwama, Ethn. for Pungupungu 84.3.1
Kwa 3.9 pp. 111, 283
Kwaami, Ethn. for Kwami 7.2.4
Kwaio 31.3.1
Kwakiutl, Ethn. for Kwakw'ala 109.1
Kwakw'ala 109.1 pp. 35, 67
Kwamera 31.3.1
Kwami 7.2.4
Kwangali 3.3.1
Kwaza 180
Kwazá 180
Kwegu, Ethn. for Koegu 6.3.8
Kwerba 39
Kwini, Ethn. for Gunin 84.21
Kwoma 52.1
Kxoe 1.1 p. 526
Kyaka 57.9
Kyuquot 109.2
Laal 4 p. 138
Lacandón 132
Lachi 28.1
Ladakhi 26.2.3 pp. 147, 298, 463
Ladin 9.9
Ladino 9.9
Lafofa 3.7
Lagwan 7.2.1
Laha 28.1
Lahu 26.2.4 pp. 450, 451, 573
Lai 26.2.7 pp. 94, 95, 358, 410, 411, 438, 442, 443
Lak 20.1.2 pp. 35, 119, 127, 131, 190, 202, 438
Lak-Dargwa 20.1.2
Lak, Ethn. for Siar 31.3.1
Lake Miwok, see Miwok (Lake) 113.7.2
Lakes Plain 38
Lakher, see Chin (Mara) 26.2.7
Lakhota 104 pp. 43, 46, 86, 154, 303, 306, 548
Lakka Koh, see Koh (Lakka) 3.1
Lakkia 28.1 pp. 35, 83
Lakota, see Lakhota 104
Lala, see Ala'ala 31.3.1
Lalo 26.2.4 p. 451
Lamaholot 31.2
Lamang 7.2.1
Lamani 9.7 p. 411
Lamba 3.3.1
Lambadi, Ethn. for Lamani 9.7
Lamé 7.2.3
Lamnso' 3.3.1
Lampung 31.6.10 p. 115
Lamu-Lamu 84.17
Langi 3.3.1
Lango 6.3.5 pp. 46, 78, 91, 94, 147, 174, 230, 282, 378, 418, 438, 506
Langue des Signes Française 207 pp. 558, 559, 561
Langue des Signes Québecoise 207 pp. 558, 559, 561
Languedocien, Ethn., see Occitan 9.9
Lao 28.2
Lappish, see Saami (Northern) 10.1.1
Laragia 84.10
Laragiyan 84.10
Lardil 84.17
Larike 31.2 p. 418
Larike-Wakasihu, Ethn. for Larike 31.2
Latin (classical) pp. 51, 66, 86, 87, 130, 165, 191, 195, 331, 467, 490, 491, 568
Latvian 9.3 pp. 207, 286, 287, 334, 398, 463
Lau 31.3.1
Lauje 31.6.9
Laven, Ethn. for Loven 29.1.2
Lavukaleve 82 pp. 158, 163, 183, 302, 438
Laz 21
Lealao Chinantec, see Chinantec (Lealao) 124.3
Lebanese Arabic, see Arabic (Lebanese) 7.6
Lebeo 3.3.1
Ledo, see Kaili 31.6.9
Lega 3.3.1 pp. 315, 319
Lega-Mwenga, Ethn., see Lega 3.3.1
Lega-Shabunda, Ethn., see Lega 3.3.1
Lekeitio Basque, see Basque (Lekeitio) 8
Lele 7.2.2
Lelemi 3.9
Lenakel 31.3.1
Lendu 6.2.3 p. 343
Lengua 197
Lengua de Señas Argentina 207 pp. 558, 559
Lengua de Señas Española 207 pp. 558, 561, 564, 565
Lepcha 26.2.8 p. 450
Lese 6.2.5
Lesser Antillean Creole French, Ethn. for Lesser Antillean French Creole 206
Lesser Antillean Creole French, Ethn., see Guadeloupe Creole 206
Lesser Antillean Creole French, Ethn., see Martinique Creole 206
Lesser Antillean French Creole 206
Leti 31.2
Lewo 31.3.1
Lezgi, Ethn. for Lezgian 20.1.3
Lezgian 20.1.3 pp. 50, 143, 158, 178, 250, 299, 314, 318, 322, 346, 459, 462, 463
Lezgic 20.1.3
Lhomi 26.2.3
Ligurian, Ethn., see Italian (Genoa) 9.9
Lillooet 110.3 pp. 234, 250, 422
Limbu 26.2.3 pp. 158, 286, 407
Limburg Dutch, see Dutch (Limburg) 9.5
Limilngan 84.11
Linda 3.1
Lingala 3.3.1 pp. 286, 295
Lingombe, see Ngombe 3.3.1
Língua de Sinais Brasileira 207 pp. 558, 559
Língua Gestual Portuguesa 207 p. 558
Lingua Italiana dei Segni 207 pp. 558, 559, 561, 564
Liptako Ful, see Ful (Liptako) 3.2.1
Lishán Didán, Ethn. for Neo-Aramaic (Persian Azerbaijan) 7.6
Lishanid Noshan, Ethn. for Neo-Aramaic (Arbel Jewish) 7.6
Lisu 26.2.4
Lithuanian 9.3 pp. 170, 183, 431, 434, 523
Lithuanian Yiddish, see Yiddish (Lithuanian) 9.5
Liv 10.1.1
Lobi 3.5
Lobiri, see Lobi 3.5
Lodz Yiddish, see Yiddish (Lodz) 9.5
Logbara, see Lugbara 6.2.6
Logo, Ethn. for Logoti 6.2.6
Logone, see Lagwan 7.2.1
Logoti 6.2.6
Loko 3.10.2
Lokono Dian, see Arawak 174
Lolovoli Northeast Ambae, see Ambae (Lolovoli Northeast) 31.3.1
Loma 3.10.2
Lomongo, see Mongo 3.3.1
Lomtomba, see Ntomba 3.3.1
Londo 3.3.1
Longa, see Amara 31.3.1
Longgu 31.3.1 p. 410
Loniu 31.3.1 pp. 43, 154, 234, 235, 282, 351,
Lonwolwol 31.3.1 pp. 487, 522
Lotha 26.2.7
Lou 31.3.1
Lovari Romani, see Romani (Lovari) 9.7
Loven 29.1.2
Low German 9.5
Lower Chinook, see Chinook (Lower) 113.1
Lower Grand Valley Dani, see Dani (Lower Grand Valley) 57.6
Lower Mamberamo 36
Lower Ramu 53.3
Lower Sepik 53.4
Lower Sepik-Ramu 53
Lower Sorbian, see Sorbian (Lower) 9.10
Lower Tanana, see Tanana (Lower) 87.1
Lower Umpqua, see Siuslaw 114.3
Lozi 3.3.1
Lü 28.2
Lua 3.1
Luang, Ethn. for Wetan 31.2
Luangiua 31.3.1
Luba-Kasai, Ethn. for CiLuba 3.3.1
Luba-Kasai, Ethn., see Bena-Lulua 3.3.1
Luba-Katanga, Ethn. for Kiluba 3.3.1
Lucazi 3.3.1 p. 250
Luchazi, Ethn. for Lucazi 3.3.1
Luganda 3.3.1 pp. 475, 482
Lugbara 6.2.6 pp. 83, 86, 90, 91
Lughat al-Isharat al-Lubnaniya 207 p. 558
Luiseño 108.6 pp. 43, 115, 239, 475, 522, 548
Lule 200
Lule-Vilela 200
Lummi 110.2 p. 43
Lun Dayeh 31.6.1
Lunda 3.3.1
Lundayeh, Ethn. for Lun Dayeh 31.6.1
Lungchang 26.2.2
Luo 6.3.5 pp. 415, 478, 479
Luri 9.8
Lushai, Ethn., see Mizo 26.2.7
Lushootseed 110.2 p. 35
Lusi, Ethn., see Kaliai-Kove 31.3.1
Luvale 3.3.1
Luwo, Ethn. for Jur Luwo 6.3.5
Luxembourgeois, Ethn. for Luxemburgeois 9.5
Luxemburgeois 9.5
Luyia 3.3.1
Lwena, see Luvale 3.3.1
Lyele 3.5
Lyélé, Ethn. for Lyele 3.5
Ma 3.1 p. 455
Ma'anyan 31.6.1
Ma'di 6.2.6 p. 343
Ma'ya 31.3.2
Maa, see Maasai 6.3.5
Maale 7.5 p. 496
Maasai 6.3.5 pp. 35, 210, 490
Maasina Fulfulde, see Fulfulde (Maasina) 3.2.1
Maba 6.8 pp. 46, 210, 358
Maban 6.8
Maca 198 p. 183
Macaguán 142
Macedonian 9.10 pp. 63, 315
Machiguenga 174 p. 523
Macro-Ge 177 p. 327
Macú, see Hupda 173
Macuna 145 p. 523
Macushi 175 pp. 131, 234, 338, 371, 415
Macuxi, see Macushi 175
Mada (in Cameroon) 7.2.1
Mada (in Nigeria) 3.3.9
Mada, Ethn. for Mada (in Cameroon) 7.2.1
Mada, Ethn. for Mada (in Nigeria) 3.3.9
Madang 57.14
Madimadi 84.17 pp. 70, 71
Madngele 84.3.2
Madura, Ethn. for Madurese 31.6.10
Madurese 31.6.10 p. 262
Mae 31.3.1
Magahi 9.7
Magar 26.2.3
Magar (Eastern), Ethn. for Magar 26.2.3
Magar (Syangja) 26.2.3
Magar (Western), Ethn. for Magar (Syangja) 26.2.3
Magi 57.15
Magindanao 31.6.8
Magindanaon, Ethn. for Magindanao 31.6.8
Mah Meri 29.1.1
Mahican 89.1
Mahou, Ethn. for Mauka 3.10.2

Mai Brat, Ethn. for Maybrat 32.4
Maidu (Northeast) 113.3 pp. 546, 548
Maidu (Northwest), Ethn. for Konkow 113.3
Maiduan 113.3 pp. 548, 549
Mailu, Ethn. for Magi 57.15
Mailuan 57.15
Maipure 174
Mairasi 57.16
Mairasi-Tanah Merah 57.16
Maisin 31.3.1
Maithili 9.7 p. 523
Majang 6.3.8 pp. 374, 383
Makaa 3.3.1
Makah 109.2 p. 147
Makasar, Ethn. for Makassar 31.6.9
Makassar 31.6.9
Makhuwa, Ethn. for Makua 3.3.1
Makian (East), Ethn. for Taba 31.3.2
Maklew 62
Makonde 3.3.1
Máku 172
Makú, see Hupda 173
Makua 3.3.1
Malacca Creole 206
Malaccan Creole Portuguese, Ethn. for Malacca Creole 206
Malagasy 31.6.1 pp. 63, 94, 171, 223, 230, 378
Malakmalak 84.3.4 pp. 50, 67, 379
Malay 31.6.10 pp. 71, 162, 231, 491, 569
Malay (Kuala Lumpur) 31.6.10
Malay (Manadonese), Ethn. for Manadonese 31.6.10
Malay (Tenggarong Kutai), Ethn. for Kutai 31.6.10
Malay (Ulu Muar) 31.6.10
Malay, Ethn., see Ketapang 31.6.10
Malay, Ethn., see Malay (Kuala Lumpur) 31.6.10
Malayalam 25.4 pp. 66, 74, 179, 194, 195, 572
Malayo, Ethn. for Damana 135.1
Maldivian, Ethn. for Dhivehi 9.7
Male, Ethn. for Maale 7.5
Malecite-Passamaquoddy, Ethn. for Passamaquoddy-Maliseet 89.1
Maleu 31.3.1
Maleu-Kilenge, Ethn. for Maleu 31.3.1
Malgwa 7.2.1
Mali 78
Mali-Baining, see Mali 78
Malo, Ethn., see Tamabo 31.3.1
Maltese 7.6 pp. 467, 475, 486
Malto 25.2
Mam 132 pp. 71, 171, 438
Mam (Central), Ethn., see Mam 132
Mam (Southern) 132
Mamanwa 31.6.3
Mamasa 31.6.9
Mambila 3.3.1
Mambila (Cameroon), Ethn., see Mambila 3.3.1
Mambila (Nigeria), Ethn., see Mambila 3.3.1
Mambwe 3.3.1
Mambwe-Lungu, Ethn. for Mambwe 3.3.1
Mampruli 3.5
Mamvu 6.2.5 p. 138
Manadonese 31.6.10
Manam 31.3.1 p. 450
Manambu 52.1 p. 407
Mancagne, see Mankanya 3.2.1
Manchu 11.2 pp. 50, 523
Mandaic (Modern) 7.6
Mandaic, Ethn. for Mandaic (Modern) 7.6
Mandan 104
Mandar 31.6.9
Mandarin 26.1 pp. 10, 26, 42, 50, 87, 90, 94, 127, 146, 151, 162, 165, 194, 227, 250, 251, 307, 330, 342, 347, 350, 351, 390, 398, 427, 451, 479, 506, 530, 531, 554, 565, 568
Mande 3.10 pp. 343, 427
Mandinka 3.10.2
Mandinka (Gambian) 3.10.2
Mandinka, Ethn., see Mandinka (Gambian) 3.10.2
Mandjak, Ethn. for Manjaku 3.2.1
Manem 46
Mangap-Mbula 31.3.1
Mangarayi, Ethn. for Mangarrayi 84.12
Mangarrayi 84.12 pp. 43, 106, 114, 154, 163, 210, 318, 334, 359, 379, 438, 530
Mangbetu 6.2.4
Mangbutu-Efe 6.2.5
Manggarai 31.2 p. 470
Mangghuer 11.1 p. 523
Manikion, Ethn. for Sougb 33
Maninka 3.10.2
Maninka (Kankan), Ethn. for Maninka 3.10.2
Maninka (Western) 3.10.2
Maninkakan (Western), Ethn. for Maninka (Western) 3.10.2
Manipuri, see Meithei 26.2.7
Manjaku 3.2.1
Mankanya 3.2.1
Mano 3.10.1
Manobo (Matigsalug), Ethn., see Kulamanen 31.6.8
Manobo (Western Bukidnon) 31.6.8
Mansfeldisch German, see German (Mansfeldisch) 9.5
Mansi 10.1.2 pp. 42, 415
Mantjiltjara 84.17
Manx 9.4
Mao Naga, see Naga (Mao) 26.2.7
Maori 31.3.1 pp. 26, 107, 127, 139, 454, 479, 503, 510
Mapoyo 175
Mapuche, see Mapudungun 202
Mapudungun 202 pp. 43, 62, 74, 75, 83, 94, 147, 202, 299, 546, 548,
Maquiritari, Ethn. for Carib (De'kwana) 175
Mara 84.13 p. 323
Mara Chin, see Chin (Mara) 26.2.7
Maran 84.13
Maranao 31.6.8
Maranunggu, Ethn. for Maranungku 84.3.6
Maranungku 84.3.6 pp. 139, 526
Marathi 9.7 pp. 127, 186, 400, 401, 523
Marchha 26.2.3
Margany 84.17 p. 42
Marghi (Central), Ethn. for Margi 7.2.1
Margi 7.2.1
Mari (High), Ethn. for Mari (Hill) 10.1.1
Mari (Hill) 10.1.1 p. 67
Mari (Low), Ethn. for Mari (Meadow) 10.1.1
Mari (Meadow) 10.1.1 pp. 67, 83, 255
Maric p. 164
Maricopa 117.7 pp. 83, 94, 138, 162, 171, 222, 226, 234, 239, 298, 299, 370, 495, 496, 503, 510, 518
Marind 59
Marind 59.1
Marind Proper 59.1
Maring 57.5
Maringarr 84.3.6
Marithiel, Ethn. for Marrithiyel 84.3.6
Marma Arakanese, see Arakanese (Marma) 26.2.4
Marquesan 31.3.1
Marquesan (North), Ethn., see Marquesan 31.3.1
Marquesan (South), Ethn., see Marquesan 31.3.1
Marrithiyel 84.3.6
Marshallese 31.3.1
Martinique Creole 206
Martu Wangka 84.17
Martu Wangka, Ethn., see Mantjiltjara 84.17
Martu Wangka, Ethn., see Yulparija 84.17
Martuthunira 84.17
Martuyhunira, Ethn. for Martuthunira 84.17
Maru 26.2.4
Marwari, Ethn., see Shekhawati 9.7
Masa 7.2.3
Masakin 3.7
Masalit 6.8
Masana, Ethn. for Masa 7.2.3
Mascoian 197
Maskoy, see Lengua 197
Massachusett 89.1
Mataco, see Wichí 198
Matacoan 198
Matis 168
Matís, Ethn. for Matis 168
Matlatzinca (Atzingo), Ethn. for Ocuilteco 124.5
Matsés 168
Mattole 87.1
Matukar 31.3.1
Matumbi, Ethn. for Matuumbi 3.3.1
Matuumbi 3.3.1
Mauka 3.10.2 p. 359
Maung 84.8 pp. 78, 127, 323, 459, 462, 463
Mauritanian Fula, see Fula (Mauritanian) 3.2.1
Mauritian Creole 206
Mawchi 9.7
Mawiha, see Makonde 3.3.1
Maxakalí 177.6 p. 79
Maya (Yucatán), Ethn. for Yucatec 132
Mayali, see Bininj Gun-Wok 84.7.2
Mayan 132 pp. 199, 407, 438, 439, 442
Maybrat 32.4 pp. 198, 250, 347, 350, 351, 366, 451, 470, 495
Mayi-Yapi 84.17
Mayo 108.2 pp. 223, 522
Mayo, see Yessan-Mayo 52.4
Mazahua 124.5
Mazahua (Central), Ethn., see Mazahua 124.5
Mazahua (Michoacán), Ethn., see Mazahua 124.5
Mazanderani 9.8
Mazatec (Chiquihuitlán) 124.7
Mazatec (Huautla) 124.7
Mazateco (Huautla), Ethn. for Mazatec (Huautla) 124.7
Mba 3.1 p. 54
Mbabaram 84.17 pp. 71, 164
Mbalanhu 3.3.1
Mbara 7.2.1
Mbay 6.2.1 pp. 147, 400
Mbe' 3.3.1
Mbere 3.3.1
Mbili 3.3.1
Mbodomo 3.1
Mbole 3.3.1
Mbugu 7.3.4
Mbula, Ethn. for Mangap-Mbula 31.3.1
Mbum 3.1
Me'en 6.3.8
Meadow Mari, see Mari (Meadow) 10.1.1
Meah, see Meyah 33
Mecayapan Isthmus Nahuatl, see Nahuatl (Mecayapan Isthmus) 108.1
Mehek 52.4
Mehri 7.6
Mehri, Ethn., see Bathari 7.6
Meitei, Ethn. for Meithei 26.2.7
Meithei 26.2.7 pp. 43, 158, 573
Mek 57.17
Mekem, Ethn. for Mekens 176.5
Mekens 176.5
Mekeo 31.3.1
Melanau 31.6.1
Melayu Betawi 206
Mende 3.10.2
Menomini 89.1 p. 491
Mentawai 31.6.10
Mentuh Tapuh 31.6.1
Menya 57.1
Meriam, Ethn. for Meryam Mir 69
Meryam Mir 69
Mesa Grande Diegueño, see Diegueño (Mesa Grande) 117.7
Meso-Philippine 31.6.3
Mesquakie, Ethn. for Fox 89.1
Metzontla Popoloca, see Popoloca (Metzontla) 124.7
Meyah 33
Mezquital Otomí, see Otomí (Mezquital) 124.5
Miami, Ethn. for Illinois 89.1
Michif 89.1
Michoacan Nahuatl, see Nahuatl (Michoacán) 108.1
Micmac 89.1
Middle Atlas Berber, see Berber (Middle Atlas) 7.1
Middle Sepik 52.1
Midob 6.3.6
Mien 27
Migaama, Ethn. for Migama 7.2.2
Migama 7.2.2
Miisiirii 6.3.9
Miju 26.2.9
Mikarew 53.3
Mikarew 53.5
Mikasuki 94
Mikir 26.2.7
Milang 26.2.9 p. 358
Milpa Alta Nahuatl, see Nahuatl (Milpa Alta) 108.1
Min Nan Chinese, Ethn., see Hokkien 26.1 p. 554
Minangkabau 31.6.10 pp. 183, 226, 227, 246, 251, 572

Minaveha 31.3.1
Mingrelian 21 p. 318
Minica Huitoto, see Huitoto (Minica) 148.2
Minnesota Ojibwe, see Ojibwe (Minnesota) 89.1
Minor Mlabri, see Mlabri (Minor) 29.1.8
Miri, Ethn. for Mising 26.2.9
Mirish 26.2.9
Miriwung 84.4
Misantla Totonac, see Totonac (Misantla) 126
Mishar Tatar, see Tatar (Mishar) 11.3
Mishmi, see Digaro 26.2.9
Mising 26.2.9 p. 354
Miskito 134 pp. 482, 548
Mískito Coast English Creole 206
Mískito, Ethn. for Miskito 134
Misumalpan 134
Mitla Zapotec, see Zapotec (Mitla) 124.8
Mituku 3.3.1
Miwok 113.7.2
Miwok (Bodega) 113.7.2
Miwok (Central Sierra) 113.7.2
Miwok (Coast), Ethn. for Miwok (Bodega) 113.7.2
Miwok (Lake) 113.7.2
Miwok (Northern Sierra) 113.7.2 p. 526
Miwok (Plains) 113.7.2 p. 526
Miwok (Southern Sierra) 113.7.2 pp. 102, 163, 548, 549
Mixe p. 548
Mixe (Coatlán) 130
Mixe (Tlahuitoltepec) 130
Mixe (Totontepec) 130
Mixe-Zoque 130
Mixtec pp. 83, 186
Mixtec (Alacatlatzala) 124.4
Mixtec (Atatlahuca) 124.4
Mixtec (Ayutla) 124.4
Mixtec (Chalcatongo) 124.4 pp. 43, 127, 139, 147, 195, 255, 282, 322, 351, 386, 411, 467, 471, 496
Mixtec (Chayuco) 124.4
Mixtec (Coatzospan) 124.4
Mixtec (Jamiltepec) 124.4
Mixtec (Jicaltepec) 124.4
Mixtec (Molinos) 124.4
Mixtec (Ocotepec) 124.4 p. 282
Mixtec (Peñoles) 124.4
Mixtec (San Juan Colorado) 124.4
Mixtec (San Miguel el Grande) 124.4
Mixtec (Silacayoapan) 124.4
Mixtec (Yosondúa) 124.4
Mixtecan 124.4 p. 283
Mixteco (Alacatlatzala), Ethn. for Mixtec (Alacatlatzala) 124.4
Mixteco (Atatláhuca), Ethn. for Mixtec (Atatlahuca) 124.4
Mixteco (Ayutla), Ethn. for Mixtec (Ayutla) 124.4
Mixteco (Chayuco), Ethn. for Mixtec (Chayuco) 124.4
Mixteco (Coatzospan), Ethn. for Mixtec (Coatzospan) 124.4
Mixteco (Jamiltepec), Ethn. for Mixtec (Jamiltepec) 124.4
Mixteco (Ocotepec), Ethn. for Mixtec (Ocotepec) 124.4
Mixteco (Peñoles), Ethn. for Mixtec (Peñoles) 124.4
Mixteco (Pinotepa Nacional), Ethn. for Mixtec (Jicaltepec) 124.4
Mixteco (San Juan Colorado), Ethn. for Mixtec (San Juan Colorado) 124.4
Mixteco (San Miguel el Grande), Ethn. for Mixtec (San Miguel el Grande) 124.4
Mixteco (San Miguel el Grande), Ethn., see Mixtec (Chalcatongo) 124.4
Mixteco (San Miguel el Grande), Ethn., see Mixtec (Molinos) 124.4
Mixteco (Silacayoapan), Ethn. for Mixtec (Silacayoapan) 124.4
Mixteco (Yosondúa), Ethn. for Mixtec (Yosondúa) 124.4
Mixtepec Zapotec, see Zapotec (Mixtepec) 124.8
Miya 7.2.4 pp. 331, 383
Mizo 26.2.7
Mlabri (Minor) 29.1.8
Mlabri, Ethn., see Mlabri (Minor) 29.1.8
Mnong (Eastern) 29.1.2
Moca 7.5
Mochica 166
Mocoví 199
Modern Ashenazic Hebrew, see Hebrew (Modern Ashenazic) 7.6
Modern Chaldean, see Chaldean (Modern) 7.6
Modern Greek, see Greek (Modern) 9.6
Modern Hebrew, see Hebrew (Modern) 7.6
Modern Literary Tibetan, see Tibetan (Modern Literary) 26.2.3
Modern Mandaic, see Mandaic (Modern) 7.6
Modern Standard Arabic, see Arabic (Modern Standard) 7.6
Mödö, see Jur Mödö 6.2.1
Mofu-Gudur 7.2.1
Moghol 11.1 p. 523
Mogholi, Ethn. for Moghol 11.1
Mohave, Ethn. for Mojave 117.7
Mohawk 91.1 pp. 111, 526
Mohegan-Montauk-Narragansett, Ethn., see Mahican 89.1
Mojave 117.7
Mokilese 31.3.1 pp. 110, 114, 335, 482
Mokilko 7.2.2
Mokpwe, Ethn. for Bakueri 3.3.1
Moksha Mordvin, see Mordvin (Moksha) 10.1.1
Moksha, Ethn. for Mordvin (Moksha) 10.1.1
Molala 113.4
Moldavian 9.9
Molinos Mixtec, see Mixtec (Molinos) 124.4
Mombum 61
Mon 29.1.6
Mon-Khmer 29.1 pp. 315, 573
Monde 176.2
Mondunga 3.1
Mongo 3.3.1
Mongo (Nkundu), Ethn. for Mongo 3.3.1
Mongol (Khamnigan) 11.1
Mongolian (Halh), Ethn. for Khalkha 11.1
Mongolian (Peripheral), Ethn., see Ordos 11.1
Mongolian (written) p. 474
Mongolic 11.1 pp. 147, 548
Mongondow 31.6.9
Monguor, see Mangghuer 11.1
Moni 57.20
Monic 29.1.6
Mono (in United States) 108.5 pp. 138, 374
Mono-Alu 31.3.1
Mono, Ethn. for Mono (in United States) 108.5
Mono, Ethn. for Mono-Alu 31.3.1
Monpa, see Tshangla 26.2.3
Montagnais 89.1
Monumbo 55
Mooré 3.5
Mòoré, Ethn. for Mooré 3.5
Mopan 132
Mopán Maya, Ethn. for Mopan 132
Mor 31.3.2
Moraori 63
Mordvin (Erzya) 10.1.1 p. 415
Mordvin (Moksha) 10.1.1 p. 415
Moré, see Mooré 3.5
Morehead and Upper Maro Rivers 63
Morisyen, Ethn. for Mauritian Creole 206
Moro 3.7 p. 83
Moroccan Arabic, see Arabic (Moroccan) 7.6
Morori, Ethn. for Moraori 63
Moru 6.2.6 p. 339
Moru-Ma'di 6.2.6 p. 339
Morwap 43
Mosetén 190 p. 230
Mosetenan 190
Motilon 135.5
Motilón (Chibchan) 135.5
Motilón, Ethn. for Motilón (Chibchan) 135.5
Motilone (Carib), see Yukpa 175
Motlav, Ethn. for Mwotlap 31.3.1
Motu 31.3.1 p. 487
Motuna 81 p. 427
Mountain Koiali, see Koiali (Mountain) 57.12
Movima 189
Mpakwithi, see Anguthimri 84.17
Mparntwe Arrernte, see Arrernte (Mparntwe) 84.17
Mpongwe 3.3.1
Mpur 32.2
Mudbura, Ethn. for Mudburra 84.17
Mudburra 84.17 p. 573
Muduapa, Ethn., see Bali-Vitu 31.3.1
Mugil 57.14
Muher 7.6
Muinane 148.1
Muinane Huitoto, see Huitoto (Muinane) 148.2
Muisca 135.2
Mukulu, Ethn. for Mokilko 7.2.2
Mulao 28.2 pp. 175, 451
Mullukmulluk, Ethn. for Malakmalak 84.3.4
Mulwi, see Musgu 7.2.1
Mumuye 3.1
Mun 27
Muna 31.6.9 pp. 334, 507
Munda 29.2 pp. 43, 163, 299
Mundang 3.1
Mundani 3.3.1
Mundari 29.2 pp. 131, 146, 147, 179, 183, 490
Mundari, Ethn., see Bhumij 29.2
Mündü 3.1
Munduruku 176.3 pp. 523, 526
Mundurukú 176.3 pp. 523, 526
Mungaka 3.3.1
Munsee 89.1
Munster Irish, see Irish (Munster) 9.4
Muong 29.1.9
Mupun 7.2.4 pp. 326, 327, 455
Mura 183
Múra-Pirahã, Ethn. for Pirahã 183
Murik 53.4 p. 67
Murinypata, see Murrinh-Patha 84.3.3
Murle 6.3.8 pp. 22, 34, 35, 146, 400
Murrinh-Patha 84.3.3 pp. 367, 526
Mursi 6.3.8
Murui Huitoto, see Huitoto (Murui) 148.2
Muruwari 84.17
Musgu 7.2.1 p. 454
Muskogean 94
Muskogee, Ethn. for Creek 94
Muskogee, Ethn., see Seminole 94
Muslim Tat, see Tat (Muslim) 9.8
Musom 31.3.1
Mussau 31.3.1
Mussau-Emira, Ethn. for Mussau 31.3.1
Mutsun 113.7.1
Muyuw 31.3.1
Mwaghavul, Ethn. for Mupun 7.2.4
Mwera 3.3.1
Mwotlap 31.3.1
Myene, Ethn., see Mpongwe 3.3.1
Mzab Berber, see Berber (Mzab) 7.1
Mzab, see Berber (Mzab) 7.1
Na-Dene 87 pp. 11, 22, 35, 78
Naasioi, Ethn. for Nasioi 81
Nabak 57.10 p. 43
Nadëb 173 pp. 330, 331
Nafaanra 3.5
Naga (Angami), Ethn., see Angami 26.2.7
Naga (Ao), Ethn. for Ao 26.2.7
Naga (Chang), Ethn. for Chang 26.2.2
Naga (Chokri), Ethn., see Angami 26.2.7
Naga (Kabui), Ethn. for Kabui 26.2.7
Naga (Lotha), Ethn. for Lotha 26.2.7
Naga (Mao) 26.2.7 p. 298
Naga (Nocte), Ethn. for Nocte 26.2.2
Naga (Sumi), Ethn. for Sema 26.2.7
Naga (Tangkhul) 26.2.7
Naga (Tarao), Ethn. for Tarao 26.2.7
Naga (Tase), Ethn., see Jugli 26.2.2
Naga (Tase), Ethn., see Lungchang 26.2.2
Naga (Zeme) 26.2.7
Naga Pidgin 206
Nagatman 49 pp. 138, 342
Nahali 23
Nahuatl pp. 26, 186
Nahuatl (Central) 108.1
Náhuatl (Guerrero), Ethn., see Nahuatl (Xalitla) 108.1
Náhuatl (Huasteca Este), Ethn., see Nahuatl (Huasteca) 108.1
Náhuatl (Huasteca Oeste), Ethn., see Nahuatl (Huasteca) 108.1
Nahuatl (Huasteca) 108.1 p. 354
Nahuatl (Huauchinango) 108.1 p. 522

Nahuatl (Istmo-Mecayapan), Ethn. for Nahuatl (Mecayapan Isthmus) 108.1
Náhuatl (Istmo-Pajapan), Ethn. for Nahuatl (Pajapan) 108.1
Nahuatl (Mecayapan Isthmus) 108.1
Nahuatl (Michoacán) 108.1 p. 334
Náhuatl (Michoacán), Ethn. for Nahuatl (Michoacán) 108.1
Nahuatl (Milpa Alta) 108.1
Nahuatl (North Puebla) 108.1 p. 15
Nahuatl (Pajapan) 108.1 p. 522
Nahuatl (Pochutla) 108.1
Náhuatl (Puebla Norte), Ethn. for Nahuatl (North Puebla) 108.1
Nahuatl (Puebla Sierra), Ethn. for Nahuatl (Sierra de Zacapoaxtla) 108.1
Nahuatl (Sierra de Zacapoaxtla) 108.1 p. 522
Nahuatl (Tetelcingo) 108.1 pp. 370, 522
Náhuatl (Tetelcingo), Ethn. for Nahuatl (Tetelcingo) 108.1
Nahuatl (Xalitla) 108.1 p. 522
Nakanai 31.3.1
Nakara, Ethn. for Nakkara 84.14
Nakh 20.2
Nakh-Daghestanian 20 pp. 35, 83, 119, 127, 299, 438
Nakkara 84.14
Nalik 31.3.1
Nama, see Khoekhoe 1.1
Nambakaengö 83 p. 50
Nambas (Big) 31.3.1 p. 487
Nambikuára 179
Nambikuára (Southern), Ethn. for Nambikuára 179
Nambikuaran 179
Nambiquara, see Nambikuára 179
Namia 52.6
Nanai 11.2 pp. 42, 546, 548
Nancowry 29.1.7
Nande 3.3.1
Nandi 6.3.5 p. 210
Nandi (Bantu), Ethn. for Nande 3.3.1
Nangikurrunggurr, Ethn. for Ngankikurungkurr 84.3.5
Nangikurrunggurr, Ethn., see Ngan'gityemerri 84.3.5
Nankina 57.10
Nanticoke 89.1
Nanumea 31.3.1
Napoletano-Calabrese, Ethn., see Italian (Napolitanian) 9.9
Napolitanian Italian, see Italian (Napolitanian) 9.9
Napu 31.6.9
Nar Phu, Ethn. for Nar-Phu 26.2.3
Nar-Phu 26.2.3
Nara 6.3.4
Nara (in Ethiopia) 6.3.4
Nara, Ethn. for Ala'ala 31.3.1
Nara, Ethn. for Nara (in Ethiopia) 6.3.4
Narinjari, see Ngarinyeri 84.17
Naro, Ethn. for Nharo 1.1
Narom 31.6.1
Narrinyeri, Ethn. for Ngarinyeri 84.17
Nasioi 81
Naskapi 89.1
Natchez 96 p. 526
Nateni 3.5
Natick, see Massachusett 89.1
Naukan Yupik, see Yupik (Naukan) 86
Nauruan 31.3.1
Navajo 87.1 pp. 15, 35, 58, 171, 238, 327, 523, 526
Naxi 26.2.10
Ndebele (in South Africa) 3.3.1
Ndjébbana 84.15 pp. 183, 526
Ndogo 3.1
Ndonga 3.3.1
Ndumu 3.3.1
Ndunga, Ethn. for Mondunga 3.1
Ndut 3.2.1
Ndyuka 206 pp. 250, 370
Nederlandse Gebarentaal 207 pp. 558, 559, 565
Negeri Sembilan Malay, Ethn. for Malay (Ulu Muar) 31.6.10
Negev Arabic, see Arabic (Negev) 7.6
Negidal 11.2 p. 42
Nehan 31.3.1
Nelemwa 31.3.1 p. 466
Nend 57.14
Nenets 10.2 pp. 42, 83, 91, 147, 415, 548
Nenets (Tundra) 10.2
Nenets, Ethn., see Nenets (Tundra) 10.2
Nengone 31.3.1
Neo-Aramaic (Amadiya) 7.6
Neo-Aramaic (Arbel Jewish) 7.6 p. 247
Neo-Aramaic (Assyrian) 7.6
Neo-Aramaic (Persian Azerbaijan) 7.6
Nepali 9.7 p. 299
Nera, see Nara (in Ethiopia) 6.3.4
Nevome 108.8
New Zealand Sign Language 207 pp. 558, 561
Newari (Dolakha) 26.2.3
Newari (Kathmandu) 26.2.3
Newari, Ethn., see Newari (Dolakha) 26.2.3
Newari, Ethn., see Newari (Kathmandu) 26.2.3
Nez Perce 113.5 pp. 114, 242, 400, 438, 526, 546, 573
Ngaanyatjarra 84.17
Ngäbere 135.3
Ngad'a 31.2
Ngadjumaja 84.17 p. 526
Ngadjunmaya, Ethn. for Ngadjumaja 84.17
Ngaju, see Oloh Mangtangai 31.6.1
Ngaju, see Pulopetak 31.6.1
Ngala p. 182
Ngalakan 84.7.3 p. 346
Ngalkbun 84.7.2 p. 75
Ngambay 6.2.1
Ngan'gityemerri 84.3.5 pp. 126, 127
Nganasan 10.2 pp. 42, 50, 83, 150, 299
Ngandi 84.7.4 p. 338
Ngankikurungkurr 84.3.5 p. 198
Ngarinman, Ethn. for Ngarinyman 84.17
Ngarinman, Ethn., see Bilinara 84.17
Ngarinyeri 84.17
Ngarinyin, Ethn. for Ungarinjin 84.21
Ngarinyman 84.17
Ngarluma 84.17
Ngas, Ethn. for Angas 7.2.4
Ngawun 84.17 pp. 164, 522
Ngbaka 3.1
Ngbandi 3.1
Ngbandi (Northern), Ethn., see Ngbandi 3.1
Ngbandi (Southern), Ethn., see Ngbandi 3.1
Ngelima, Ethn. for Lebeo 3.3.1
Ngemba 3.3.1
Ngile, Ethn. for Masakin 3.7
Ngiti 6.2.3 pp. 83, 138, 164, 195, 343, 359, 374
Ngiyambaa 84.17 pp. 78, 83, 114, 155, 165, 367, 430, 431, 526
Ngizim 7.2.4 pp. 35, 526, 527
Ngkâlmpw Kanum, see Kanum (Ngkâlmpw) 63
Ngombe 3.3.1
Ngombe (Bantu), Ethn. for Ngombe 3.3.1
Ngombe (Bantu), Ethn., see Doko 3.3.1
Ngoni 3.3.1
Ngukurr Kriol, see Kriol (Ngukurr) 206
Ngumba, Ethn. for Bujeba 3.3.1
Nguna 31.3.1
Ngura, Ethn., see Wangkumara 84.17
Nhanda 84.17 p. 262
Nhang, Ethn. for Yay 28.2
Nharo 1.1
Nias 31.6.10 pp. 114, 210, 330, 362, 363, 398, 411
Nicobarese 29.1.7 pp. 127, 131, 222, 223
Nicobarese (Car) 29.1.7
Nicobarese (Central), Ethn., see Nancowry 29.1.7
Nicobarese (Central), Ethn., see Nicobarese 29.1.7
Nicobarese, see Nancowry 29.1.7
Niellim, Ethn. for Lua 3.1
Niger-Congo 3 pp. 11, 15, 27, 35, 58, 82, 83, 86, 122, 127, 131, 135, 139, 183, 194, 269, 283, 287, 291, 295, 343, 395, 415, 427, 435, 447, 548
Nigerian Fula, see Fula (Nigerian) 3.2.1
Nigerian Pidgin 206
Nihali, Ethn. for Nahali 23
Nihon Shuwa (Japanese Sign Language) 207 pp. 558, 559, 561, 565
Nilo-Saharan 6 pp. 15, 27, 35, 83, 86, 122, 127, 175, 183, 203, 269, 283, 323, 327a, 335, 400, 423, 435, 447
Nilotic 6.3.5 pp. 86, 210, 335, 339
Nimboran 41
Ninam, Ethn. for Shiriana 154
Ningil 51.3
Nisenan 113.3
Nisga'a, Ethn. for Nisgha 113.6
Nisgha 113.6 p. 546
Nishi 26.2.9 p. 358
Nisi, Ethn. for Nishi 26.2.9
Nissan, see Nehan 31.3.1
Nitinaht 109.2
Niuafo'ou 31.3.1
Niue, Ethn. for Niuean 31.3.1
Niuean 31.3.1 p. 374
Nivkh 15 pp. 42, 122, 147, 163, 190, 226, 235, 287, 362, 386, 407, 415, 463
Nivkh (South Sakhalin) 15
Nkem 3.3.1
Nkem-Nkum, Ethn., see Nkem 3.3.1
Nkonya 3.9
Nkore-Kiga 3.3.1 pp. 258, 287, 382, 471, 523
Nobiin 6.3.6
Nocte 26.2.2 p. 414
Nogai, Ethn. for Noghay 11.3
Nogai, Ethn., see Noghay (Karagash) 11.3
Nogai, Ethn., see Tatar-Noghay (Alabugat) 11.3
Nogai, Ethn., see Yurt Tatar 11.3
Noghay 11.3 p. 523
Noghay (Karagash) 11.3
Nomad 65
Nomatsiguenga 174 p. 523
Noni 3.3.1 p. 378
Noon 3.2.1
Noone, Ethn. for Noni 3.3.1
Nootka, Ethn. for Nuuchahnulth 109.2
Nootka, Ethn., see Kyuquot 109.2
Nootka, Ethn., see Nitinaht 109.2
Nopala Chatino, see Chatino (Nopala) 124.8
Norsk Tegnspråk 207 p. 558
North Frisian, see Frisian (North) 9.5
North Halmaheran 32.3
North Levantine Spoken Arabic, see Arabic (North Levantine Spoken) 7.6
North Puebla Nahuatl, see Nahuatl (North Puebla) 108.1
North Russian Romani, see Romani (North Russian) 9.7
North-Central Bird's Head 32.4
Northeast Maidu, see Maidu (Northeast) 113.3
Northern Atlantic 3.2.1
Northern Barasano, see Barasano (Northern) 145
Northern Central America Creole, Ethn., see Mískito Coast English Creole 206
Northern Chukotko-Kamchatkan 14.1
Northern Daly 84.3.4
Northern Dravidian 25.2
Northern Haida, see Haida (Northern) 88
Northern High Navarrese Basque, see Basque (Northern High Navarrese) 8
Northern Iroquoian 91.1
Northern Khoisan 1.2
Northern Nuni, see Nuni (Northern) 3.5
Northern Paiute, see Paiute (Northern) 108.5
Northern Philippines 31.6.4
Northern Pomo, see Pomo (Northern) 117.3
Northern Roglai, see Roglai (Northern) 31.6.10
Northern Saami, see Saami (Northern) 10.1.1
Northern Sahaptin, see Sahaptin (Northern) 113.5
Northern Sierra Miwok, see Miwok (Northern Sierra) 113.7.2
Northern Sotho, see Sotho (Northern) 3.3.1
Northern Tepehuan, see Tepehuan (Northern) 108.8
Northern Tiwa, see Tiwa (Northern) 105
Northern Tutchone, see Tutchone (Northern) 87.1
Northern Uzbek, see Uzbek (Northern) 11.3
Northern Wakashan 109.1
Northern Zhuang, see Zhuang (Northern) 28.2
Northern Zulu, see Zulu (Northern) 3.3.1
Northwest Caucasian 19 pp. 15, 83, 155, 235, 258, 299, 479
Norton Sound Yup'ik, see Yup'ik (Norton Sound) 86
Norwegian 9.5 pp. 50, 58, 195
Norwegian (Bokmaal), Ethn. for Norwegian 9.5
Norwegian Sign Language, Ethn. for Norsk Tegnspråk 207

Nsenga 3.3.1
Ntomba 3.3.1
Nuaulu 31.2 pp. 110, 111
Nuaulu (North), Ethn., see Nuaulu 31.2
Nuaulu (South), Ethn., see Nuaulu 31.2
Nubi 206
Nubian 6.3.6
Nubian (Dongolese) 6.3.6 pp. 67, 122, 159, 195
Nubian (Hill) 6.3.6
Nubian (Kunuz) 6.3.6 pp. 346, 394
Nubian (Mahas), see Nobiin 6.3.6
Nubian (Meidob), see Midob 6.3.6
Nuer 6.3.5 pp. 138, 210, 339, 343
Nugunu, Ethn. for Gunu 3.3.1
Nukak 146
Nukak Makú, Ethn. for Nukak 146
Nukuoro 31.3.1
Numfor, see Biak 31.3.2
Numic 108.5 p. 523
Nung (in Vietnam) 28.2
Nung, Ethn. for Anong 26.2.11
Nung, Ethn. for Nung (in Vietnam) 28.2
Nunggubuyu 84.7.5 pp. 183, 286, 294, 303, 323, 330, 526
Nungish 26.2.11
Nuni (Northern) 3.5
Nupe 3.3.8
Nupe-Nupe-Tako, Ethn. for Nupe 3.3.8
Nupoid 3.3.8
Nuri, see Domari 9.7
Nusu 26.2.4 p. 363
Nuuchahnulth 109.2 pp. 35, 39, 83
Nyah Kur (Tha Pong) 29.1.6
Nyahkur, Ethn., see Nyah Kur (Tha Pong) 29.1.6
Nyâlayu, Ethn. for Nyelâyu 31.3.1
Nyambo 3.3.1
Nyamkad 26.2.3
Nyamwezi 3.3.1 p. 523
Nyang'i, Ethn. for Nyangi 6.3.3
Nyanga 3.3.1
Nyangi 6.3.3
Nyangumarda 84.17
Nyangumarta, Ethn. for Nyangumarda 84.17
Nyanja, Ethn. for Chichewa 3.3.1
Nyankore, Ethn. for Runyankore 3.3.1
Nyaturu, Ethn. for Rimi 3.3.1
Nyawaygi 84.17
Nyelâyu 31.3.1 p. 226
Nyigina 84.16
Nyiha 3.3.1
Nyimang 6.3.7
Nyoro, Ethn. for Runyoro-Rutooro 3.3.1
Nyulnyul 84.16
Nyulnyulan 84.16
Nyunga, Ethn. for Nyungar 84.17
Nyungar 84.17
Nzakara 3.1
O'odham 108.8 pp. 522, 548, 573
O'odham, Ethn., see Nevome 108.8
Obokuitai 38
Obolo 3.3.2 pp. 362, 363, 502, 503
Ocaina 148.2
Occitan 9.9
Oceanic 31.3.1
Ocotepec Mixtec, see Mixtec (Ocotepec) 124.4
Ocuilteco 124.5 p. 471
Ogbia 3.3.2
Ogbronuagum 3.3.2 p. 523
Oi 29.1.2
Oirat 11.1
Ojibwa (Eastern) 89.1 pp. 14, 127, 131, 134, 154, 354, 438, 569
Ojibwa (Severn) 89.1
Ojibwe (Minnesota) 89.1 p. 526
Ok 57.18
Okanagan 110.3
Oksapmin 50 p. 43
Oku 3.3.1
Olo 51.3 p. 394
Oloh Mangtangai 31.6.1
Olutec 130
Omagua 176.6 p. 523
Omaha 104
Omaha-Ponca, Ethn., see Omaha 104
Ömie 57.12
Omotic 7.5
Ona, Ethn. for Selknam 204.1
Oñati Basque, see Basque (Oñati) 8
One 51.4
Oneida 91.1 pp. 58, 78, 79, 203, 322, 323, 526
Onge 30.2 p. 43
Önge, Ethn. for Onge 30.2
Ongtong Java, Ethn. for Luangiua 31.3.1
Ono 57.10 p. 74
Onondaga 91.1 p. 526
Ooldea Western Desert, see Western Desert (Ooldea) 84.17
Ora (in Nigeria) 3.3.4
Ordos 11.1
Oregon Coast 114
Orejón 145 p. 523
Orig 3.7
Oriya 9.7
Oriya (Adivasi), Ethn. for Oriya (Kotia) 9.7
Oriya (Kotia) 9.7 p. 523
Ormuri 9.8
Oroch 11.2
Orok 11.2 p. 146
Orokaiva 57.4
Orokolo 72
Oromo (Boraana) 7.3.3
Oromo (Eastern), Ethn. for Oromo (Harar) 7.3.3
Oromo (Harar) 7.3.3 pp. 282, 306, 323, 378, 390, 398, 400
Oromo (Waata) 7.3.3
Oromo (West-Central) 7.3.3 p. 526
Orya 40
Osage 104
Osetin, Ethn. for Ossetic 9.8
Oshinonga, see Ndonga 3.3.1
Oshiwambo, see Ndonga 3.3.1
Ossetic 9.8 pp. 66, 467, 548
Ostschweiz German, see German (Ostschweiz) 9.5
Ostuacan Zoque, see Zoque (Ostuacan) 130
Ostyak, see Khanty 10.1.2
Òtètèla, see Tetela 3.3.1
Otjiherero, see Herero 3.3.1
Oto 104
Oto-Manguean 124 p. 323
Otomí (Ixtenco) 124.5
Otomi (Ixtenco), Ethn. for Otomí (Ixtenco) 124.5
Otomí (Mezquital) 124.5 pp. 83, 94, 195, 267
Otomi (Northwestern), Ethn., see Otomí (Santiago Mexquititlan) 124.5
Otomí (Santiago Mexquititlan) 124.5
Otomi (Sierra Oriental), Ethn. for Otomí (Sierra) 124.5
Otomí (Sierra) 124.5
Otomian 124.5
Otoro 3.7
Oy, Ethn. for Oi 29.1.2
Oyster Bay to Pitwater Tasmanian, see Tasmanian (Oyster Bay to Pitwater) 85
Pa'a 7.2.4 pp. 154, 158
Pa'anci, see Pa'a 7.2.4
Paakantyi 84.17 pp. 171, 338, 526
Paama, Ethn. for Paamese 31.3.1
Paamese 31.3.1 pp. 22, 163, 195, 198, 199, 239, 435, 487
Pacoh 29.1.3 pp. 35, 523, 526
Padoe 31.6.9
Páez 139 pp. 462 (akzent !), 502, 503
Páezan 139
Pagu 32.3
Paharia (Kumarbhag), Ethn., see Malto 25.2
Paita 31.3.1
Paite 26.2.7
Paiute (Northern) 108.5 pp. 522, 523
Paiute (Southern) 108.5
Paiwan 31.4 pp. 286, 438
Paiwanic 31.4
Pajapan Nahuatl, see Nahuatl (Pajapan) 108.1
Pakaásnovos, Ethn. for Wari' 182
Pakanha 84.17
Pakistan Sign Language, Ethn., see Indo-Pakistani Sign Language (Karachi dialect) 207
Pala, see Patpatar 31.3.1
Palaihnihan 117.2
Palantla Chinantec, see Chinantec (Palantla) 124.3
Palauan 31.6.5 pp. 183, 415, 470
Palaung 29.1.8
Palaung (Pale), Ethn., see Palaung 29.1.8
Palaung-Khmuic 29.1.8
Palestinian Arabic, see Arabic (Palestinian) 7.6
Pali p. 523
Palikur 174 p. 418
Palikúr, Ethn. for Palikur 174
Palor 3.2.1
Pama-Nyungan 84.17 pp. 99, 119, 164, 347, 411, 438
Pame 124.6
Pame (Central), Ethn., see Pame 124.6
Pamean 124.6
Pamona 31.6.9
Pampa, see Gününa Küne 204.2
Pampangan, Ethn. for Kapampangan 31.6.4
Panamint, Ethn. for Tümpisa Shoshone 108.5
Panare 175 pp. 230, 239
Pangasinan 31.6.4 pp. 114, 115
Pangwa 3.3.1
Panjabi 9.7 pp. 367, 514, 523
Panjabi (Eastern), Ethn. for Panjabi 9.7
Panoan 168 pp. 207, 327, 548
Panyjima 84.17 pp. 415, 426, 502
Panytyima, Ethn. for Panyjima 84.17
Papago, see O'odham 108.8
Papantla Totonac, see Totonac (Papantla) 126
Papiamentu 206
Parauk 29.1.8
Pare 3.3.1
Parecis 174 p. 415
Parecís, Ethn. for Parecis 174
Päri 6.3.5 pp. 334, 438
Parji (Dravidian) 25.1
Parkwa, Ethn. for Podoko 7.2.1
Parnkala, see Banggarla 84.17
Pashto 9.8 pp. 70, 569
Pashto (Central), Ethn., see Pashto 9.8
Passamaquoddy-Maliseet 89.1 pp. 131, 164, 165
Patep 31.3.1
Patpatar 31.3.1
Pattani 26.2.3
Patwin 113.8
Paulohi 31.2
Paumarí 184 pp. 62, 114, 235, 334, 400, 401, 415
Pawaia, Ethn. for Pawaian 71.1
Pawaian 71.1
Pawnee 103
Paya 135.6
Paya, see Pech 135.6
Peba-Yaguan 159
Pech 135.6
Pemon 175
Pengo 25.3
Peñoles Mixtec, see Mixtec (Peñoles) 124.4
Penutian 113 pp. 194, 548, 549
Permyak, see Komi-Permyak 10.1.1
Pero 7.2.4 p. 426
Persian 9.8 pp. 155, 190, 199, 223, 386, 400, 415, 496, 548, 569
Persian Azerbaijan Neo-Aramaic, see Neo-Aramaic (Persian Azerbaijan) 7.6
Pévé, Ethn. for Lamé 7.2.3
Phlong 26.2.6
Piapoco 174 p. 523
Piaroa 143
Picuris, see Tiwa (Northern) 105
Pidgin (Nigerian), Ethn. for Nigerian Pidgin 206
Piemontese, Ethn. for Italian (Turinese) 9.9
Pilagá 199
Pileni 31.3.1
Pima (Lower), see Pima Bajo 108.8
Pima Bajo 108.8
Pima Bajo (Chihuahua), Ethn., see Pima Bajo 108.8
Pima Bajo (Sonora), Ethn., see Pima Bajo 108.8
Pingelapese, Ethn. for Pingilapese 31.3.1
Pingilapese 31.3.1
Pintupi 84.17 p. 526
Pintupi-Luritja, Ethn. for Pintupi 84.17
Pipil 108.1 pp. 107, 138, 546, 548
Pirahã 183 pp. 47, 67, 71, 79, 94, 106, 146, 162, 470, 495, 518, 530
Piratapuyo 145 p. 523
Piro 174 p. 523
Pisa 57.3

Pitjantjatjara 84.17 pp. 147, 207, 483, 486, 526
Pitta Pitta 84.17
Plains Cree, see Cree (Plains) 89.1
Plains Miwok, see Miwok (Plains) 113.7.2
Plains-Indians Sign Language 207 pp. 558, 565
Platoid 3.3.9
Playero 142
Plei Bong-Mang Yang Bahnar, see Bahnar (Plei Bong-Mang Yang) 29.1.2
Po-Ai 31.3.1
Pochutla Nahuatl, see Nahuatl (Pochutla) 108.1
Pocomam 132
Pocomam Oriental 132
Pocomchi, see Pokomchí 132
Podoko 7.2.1 p. 195
Podopa 71.2
Pogolo, Ethn. for Pogoro 3.3.1
Pogoro 3.3.1
Pohnopeian, Ethn. for Pohnpeian 31.3.1
Pohnpeian 31.3.1 pp. 174, 239, 350, 379
Pokau, see Ala'ala 31.3.1
Poko-Rawo 45.1
Pokomam (Central), Ethn., see Pocomam 132
Pokomam (Eastern), Ethn. for Pocomam Oriental 132
Pokomchí 132
Pokomchí (Eastern), Ethn., see Pokomchí 132
Pokomchí (Western), Ethn., see Pokomchí 132
Pökoot, Ethn. for Pokot 6.3.5
Pokot 6.3.5
Polabian 9.10
Polish 9.10 pp. 63, 159, 186, 198, 434, 523
Pomo (Central) 117.3
Pomo (Eastern) 117.3 p. 179
Pomo (Northern) 117.3
Pomo (Southeastern) 117.3 p. 147
Pomoan 117.3 pp. 115, 315
Ponapean, see Pohnpeian 31.3.1
Pongu, Ethn., see Akwa 3.3.1
Popoloca (Coyotepec), Ethn., see Popoloca (San Vicente Coyotepec) 124.7
Popoloca (Metzontla) 124.7
Popoloca (Mezontla), Ethn. for Popoloca (Metzontla) 124.7
Popoloca (San Juan Atzingo) 124.7
Popoloca (San Vicente Coyotepec) 124.7
Popolocan 124.7
Popoluca (Oluta), Ethn. for Olutec 130
Popoluca (Sayula), Ethn. for Sayultec 130
Popoluca (Sierra), Ethn. for Zoque (Soteapan) 130
Poqomam, see Pocomam 132
Portuguese 9.9 pp. 22, 127, 191, 195, 554
Portuguese Sign Language, Ethn. for Língua Gestual Portuguesa 207
Potawatomi 89.1
Powhatan 89.1
Prasuni 9.7 p. 210
Príncipense 206
Principense, Ethn. for Príncipense 206
Prinmi, see Pumi 26.2.12
Provençal 9.9
Puelche 204.2
Puelche, Ethn. for Gününa Küne 204.2
Puinave 144
Pulaar, Ethn. for Fula (Senegal) 3.2.1
Pulaar, Ethn., see Fula (Mauritanian) 3.2.1
Pulopetak 31.6.1
Puluwat 31.3.1 pp. 71, 455
Puluwatese, Ethn. for Puluwat 31.3.1
Pumi 26.2.12 pp. 362, 370
Pumi (Northern), Ethn., see Pumi 26.2.12
Pumi (Southern), Ethn., see Pumi 26.2.12
Pungupungu 84.3.1
Punjabi, see Panjabi 9.7
Punu 27
Puoc, Ethn. for Ksingmul 29.1.8
Puquina 192 p. 471
Purépecha 125
Purik, Ethn. for Purki 26.2.3
Purki 26.2.3
Pwo Karen, see Karen (Pwo) 26.2.6
Q'eqchi', see K'ekchí 132
Qafar 7.3.3 pp. 127, 134
Qaqet, Ethn. for Baining 78
Qawasqar 203 pp. 30, 162
Qiang 26.2.12
Qiang (Northern), Ethn., see Qiang 26.2.12
Qiang (Southern), Ethn., see Qiang 26.2.12
Qiangic 26.2.12
Quebec Sign Language, Ethn. for Langue des Signes Québecoise 207
Quebec-Labrador Inuktitut, see Inuktitut (Quebec-Labrador) 86
Quechan 117.7
Quechua (Ancash Chiquian), Ethn., see Quechua (Ancash) 169
Quechua (Ancash) 169
Quechua (Ayacucho) 169
Quechua (Bolivian) 169
Quechua (Cajamarca) 169
Quechua (Cochabamba) 169 pp. 35, 83
Quechua (Cuzco) 169
Quechua (Ecuadorean) 169
Quechua (Huallaga Huánuco), Ethn. for Quechua (Huallaga) 169
Quechua (Huallaga) 169
Quechua (Imbabura) 169 pp. 94, 254, 375, 415, 430, 462, 471
Quechua (Junín), Ethn. for Quechua (Tarma) 169
Quechua (North Bolivian), Ethn., see Quechua (Bolivian) 169
Quechua (South Bolivian), Ethn., see Quechua (Bolivian) 169
Quechua (South Bolivian), Ethn., see Quechua (Cochabamba) 169
Quechua (Tarma) 169
Quechuan 169
Quiché 132
Quiché (Central), Ethn., see Quiché 132
Quichua (Highland, Chimborazo), Ethn., see Quechua (Ecuadorean) 169
Quichua (Highland, Imbabura), Ethn. for Quechua (Imbabura) 169
Quiegolani Zapotec, see Zapotec (Quiegolani) 124.8
Quileute 111 pp. 79, 171, 331, 523, 526
Quiotepec Chinantec, see Chinantec (Quiotepec) 124.3
Rade 31.6.10
Raga 31.3.1
Ram 52.2
Rama 135.7 pp. 43, 195, 267, 323, 454, 455, 479, 546, 548
Ramarama 176.4
Rangi, Ethn. for Langi 3.3.1
Rankin Inlet Inuktitut, see Inuktitut (Rankin Inlet) 86
Rao 53.1
Rapa Nui, Ethn. for Rapanui 31.3.1
Rapanui 31.3.1 pp. 127, 400
Rapoisi, Ethn. for Konua 80
Rashad 3.7
Ratahan, Ethn. for Toratán 31.6.9
Rawa 57.10
Rawang 26.2.11
Rawo, Ethn. for Poko-Rawo 45.1
Rayon Zoque, see Zoque (Rayon) 130
Red Lake and Pillager Chippewa, see Chippewa (Red Lake and Pillager) 89.1
Reef Islands – Santa Cruz 83
Rejang 31.6.10
Rembarnga 84.7.6
Rembarunga, Ethn. for Rembarnga 84.7.6
Remo 29.2 p. 158
Rendille 7.3.3 p. 267
Rengao 29.1.2
Rennell, Ethn. for Rennellese 31.3.1
Rennellese 31.3.1
Resígaro 174 p. 523
Retuarã 145 pp. 126, 422, 523
rGyalrong 26.2.13
Rhaeto-Romansch, see Romansch 9.9
Riantana 60
Rif Berber, see Berber (Rif) 7.1
Rif, see Berber (Rif) 7.1
Rikbaktsa 177.7
Rimi 3.3.1
Rio Grande Tewa, see Tewa (Rio Grande) 105
Ripuarian German, see German (Ripuarian) 9.5
Ritarungo, Ethn. for Ritharngu 84.17
Ritharngu 84.17
Roglai (Northern) 31.6.10
Romance 9.9 pp. 58, 63, 270, 271, 291, 315, 355, 450, 467, 573
Romani (Ajia Varvara) 9.7
Romani (Balkan), Ethn., see Romani (Ajia Varvara) 9.7
Romani (Balkan), Ethn., see Romani (Bugurdzi) 9.7
Romani (Baltic), Ethn., see Romani (North Russian) 9.7
Romani (Bugurdzi) 9.7
Romani (Burgenland) 9.7
Romani (Kalderash) 9.7
Romani (Lovari) 9.7
Romani (North Russian) 9.7
Romani (Sepecides) 9.7
Romani (Sinte), Ethn., see Romani (Burgenland) 9.7
Romani (Vlax), Ethn., see Romani (Kalderash) 9.7
Romani (Vlax), Ethn., see Romani (Lovari) 9.7
Romani (Welsh) 9.7
Romanian 9.9 pp. 271, 315, 355
Romanian, Ethn., see Moldavian 9.9
Romansch 9.9
Romansch (Scharans) 9.9
Romansch (Surmeiran) 9.9
Romansch (Sursilvan) 9.9
Romansch, Ethn., see Romansch (Scharans) 9.9
Romansch, Ethn., see Romansch (Surmeiran) 9.9
Romansch, Ethn., see Romansch (Sursilvan) 9.9
Ron 7.2.4
Roncalese Basque, see Basque (Roncalese) 8
Ronga 3.3.1
Roro 31.3.1
Rote, Ethn. for Roti 31.2
Roti 31.2
Rotokas 80 pp. 10, 79, 83
Rotuman 31.3.1
Roviana 31.3.1
Rukai 31.5
Rumanian, see Romanian 9.9
Rumu 70 pp. 390, 455
Rundi 3.3.1
Runga 6.8 p. 362
Runyankore 3.3.1
Runyoro-Rutooro 3.3.1
Rushan 9.8
Russian 9.10 pp. 10, 22, 43, 67, 87, 90, 91, 118, 119, 126, 127, 130, 131, 134, 135, 138, 146, 150, 151, 186, 190, 194, 202, 203, 211, 223, 255, 266, 268, 298, 307, 322, 323, 330, 363, 450, 459, 466, 486, 487, 496, 523, 531, 548
Russian Sign Language 207 pp. 558, 559, 560, 561
Russian-Chinese Pidgin (Birobidjan) 206
Rutul 20.1.3 pp. 83, 282
Rwanda, Ethn. for Kinyarwanda 3.3.1
Sa'a 31.3.1
Sa'ban 31.6.1
Saami (Central-South) 10.1.1
Saami (Kildin) 10.1.1
Saami (Northern) 10.1.1 p. 306
Saami (Southern), Ethn. for Saami (Central-South) 10.1.1
Saanich 110.2
Sabu, Ethn. for Sawu 31.2
Saek 28.2
Sahaptian 113.5 pp. 115, 548, 549
Sahaptin (Northern) 113.5 p. 548
Sahaptin (Umatilla) 113.5
Saharan 6.9
Sahidic, see Coptic 7.4
Saho 7.3.3
Sahu 32.3
Sakana Basque, see Basque (Sakana) 8
Sakao 31.3.1
Salar 11.3
Sáliba (in Colombia) 143
Saliba (in Papua New Guinea) 31.3.1
Sáliba, Ethn. for Sáliba (in Colombia) 143
Saliba, Ethn. for Saliba (in Papua New Guinea) 31.3.1
Sáliban 143
Salinan 121 pp. 334, 418
Salish (Samish Straits) 110.2
Salish (Southern Puget Sound) 110.2
Salish (Straits) 110.2
Salish (Straits), Ethn., see Lummi 110.2

Salish (Straits), Ethn., see Saanich 110.2
Salish (Straits), Ethn., see Salish (Samish Straits) 110.2
Salish (Straits), Ethn., see Songish 110.2
Salish (Straits), Ethn., see Sooke 110.2
Salishan 110 pp. 11, 35, 115, 175, 211, 283, 438, 442
Saliva, see Sáliba (in Colombia) 143
Salluit Inuktitut, see Inuktitut (Salluit) 86
Salt-Yui 57.5 p. 548
Sama (Balangingi) 31.6.6
Sama (Southern), Ethn., see Bajau (Semporna) 31.6.6
Sama Bajau, see Bajau 31.6.6
Sama-Bajaw 31.6.6
Samba Leko 3.1
Samish Straits Salish, see Salish (Samish Straits) 110.2
Samo, Ethn. for Nomad 65
Samoan 31.3.1 p. 330
Samoyedic 10.2 p. 42
San Andrés Tzotzil, see Tzotzil (San Andrés) 132
San Felipe Usila Chinantec, see Chinantec (San Felipe Usila) 124.3
San Juan Atzingo Popoloca, see Popoloca (San Juan Atzingo) 124.7
San Juan Colorado Mixtec, see Mixtec (San Juan Colorado) 124.4
San Juan Pueblo Tewa, see Tewa (San Juan Pueblo) 105
San Lucas Quiaviní Zapotec, see Zapotec (San Lucas Quiaviní) 124.8
San Mateo del Mar Huave, see Huave (San Mateo del Mar) 131
San Miguel el Grande Mixtec, see Mixtec (San Miguel el Grande) 124.4
San Vicente Coyotepec Popoloca, see Popoloca (San Vicente Coyotepec) 124.7
San'ani Arabic, see Arabic (San'ani) 7.6
Sandawe 1.3 p. 82
Sangir 31.6.9
Sanglechi-Ishkashimi, Ethn., see Ishkashmi 9.8
Sango 3.1 pp. 111, 155, 475
Sangu 3.3.1
Sanskrit p. 150
Santa 11.1
Santa Ana Keresan, see Keresan (Santa Ana) 106
Santa Ana Pueblo, see Keresan (Santa Ana) 106
Santa Cruz, Ethn. for Nambakaengö 83
Santali 29.2 pp. 115, 299
Santiago Mexquititlan Otomí, see Otomí (Santiago Mexquititlan) 124.5
Sanuma 154 pp. 139, 438
Sanumá, Ethn. for Sanuma 154
Sanye, Ethn. for Oromo (Waata) 7.3.3
Saposa, Ethn. for Taiof 31.3.1
Sapuan 29.1.2 p. 218
Sara-Ngambay, see Ngambay 6.2.1
Saramaccan 206
Sarcee 87.1
Sardinian 9.9 p. 572
Sardinian (Campidanese), Ethn., see Sardinian 9.9
Sare 52.3
Sarsi, Ethn. for Sarcee 87.1
Saryg Yughur 11.3
Sau, see Savi 9.7
Savi 9.7 p. 339
Savosavo 82
Sawai 31.3.2
Sawila, Ethn. for Tanglapui 57.19
Sawu 31.2
Saxon (Low), Ethn. for Low German 9.5
Sayultec 130
Scharans Romansch, see Romansch (Scharans) 9.9
Scots Gaelic, see Gaelic (Scots) 9.4
Sebei 6.3.5
Secoya 145 p. 523
Sedang 29.1.2
Sedik, see Seediq 31.1
Seediq 31.1
Selaru 31.2
Selayar 31.6.9
Selepet 57.10
Selkirk, see Tutchone (Northern) 87.1
Selknam 204.1 pp. 259, 331, 362
Selkup 10.2 pp. 50, 223, 415
Sema 26.2.7
Semai 29.1.1 p. 522
Semandang 31.6.1
Seme 3.8
Semelai 29.1.1 pp. 191, 250, 400, 573
Seminole 94
Semitic 7.6 pp. 35, 83, 122, 182, 307, 331, 335, 387, 487, 569
Semporna Bajau, see Bajau (Semporna) 31.6.6
Sena 3.3.1
Senadi 3.5
Senagi 47
Seneca 91.1 pp. 22, 26, 38, 46, 526
Senegal Fula, see Fula (Senegal) 3.2.1
Sengele 3.3.1
Senoi, see Semai 29.1.1
Senoufo (Supyire), Ethn. for Supyire 3.5
Senoufou (Cebaara), Ethn. for Senadi 3.5
Senoufou (Djimini), Ethn. for Dyimini 3.5
Sentani 42 pp. 407, 415, 463
Sepecides Romani, see Romani (Sepecides) 9.7
Sepik 52
Sepik Hill 52.3
Serbian-Croatian 9.10 pp. 179, 203, 523, 573
Serbo-Croatian, Ethn. for Serbian-Croatian 9.10
Serbo-Croatian, Ethn., see Bosnian 9.10
Seri 117.4 pp. 523, 526
Serrano 108.6
Seselwa Creole French, Ethn. for Seychelles Creole 206
Sesotho 3.3.1 pp. 231, 443, 471
Sestwana, see Tswana 3.3.1
Severn Ojibwa, see Ojibwa (Severn) 89.1
Seychelles Creole 206
Sgaw Karen, see Karen (Sgaw) 26.2.6
Shabo 6.10
Shahrudi, Ethn., see Koluri 9.8
Shakacho, Ethn. for Moca 7.5
Shambala 3.3.1 p. 154
Shan 28.2
Shapsugh Adyghe, see Adyghe (Shapsugh) 19
Sharanahua 168
Shasta 117.5 p. 526
Shatt 6.3.1
Shawnee 89.1
Shekhawati 9.7
Sherdukpen 26.2.9
Sherpa 26.2.3
Shigatse Tibetan, see Tibetan (Shigatse) 26.2.3
Shilha, see Tashlhiyt 7.1
Shilluk 6.3.5 p. 210
Shina 9.7
Shinassha 7.5
Shipibo-Conibo, Ethn. for Shipibo-Konibo 168
Shipibo-Konibo 168 pp. 226, 327, 362, 427, 458
Shira Yughur 11.1
Shiriana 154
Shokleng, see Xokleng 177.2
Shona 3.3.1 p. 523
Shona, Ethn., see Karanga 3.3.1
Shor 11.3
Shoshone 108.5 pp. 522, 523
Shoshone (Wind River) 108.5
Shoshoni, Ethn. for Shoshone 108.5
Shoshoni, Ethn., see Shoshone (Wind River) 108.5
Shuar, Ethn. for Jivaro 162
Shugnhi, Ethn., see Rushan 9.8
Shuswap 110.3 p. 283
Siamou, Ethn. for Seme 3.8
Siane 57.8
Siar 31.3.1
Siberian Yupik, see Yupik (Siberian) 86
Sie, Ethn. for Erromangan 31.3.1
Sierra de Zacapoaxtla Nahuatl, see Nahuatl (Sierra de Zacapoaxtla) 108.1
Sierra Occidental Chatino, see Chatino (Sierra Occidental) 124.8
Sierra Otomí, see Otomí (Sierra) 124.5
Sierra Totonac, see Totonac (Sierra) 126
Sign Languages 207
Sika 31.2
Sikaritai 38
Sikkimese 26.2.3
Sikuani 142
Sikule 31.6.10
Sila 6.3.1
Silacayoapan Mixtec, see Mixtec (Silacayoapan) 124.4
Silozi, see Lozi 3.3.1
Simeulue 31.6.10 p. 354
Sinasina, Ethn., see Tabare 57.5
Sinaugoro 31.3.1 p. 139
Sindhi 9.7
Sinhala 9.7 pp. 187, 486, 523
Sinhalese, see Sinhala 9.7
Sino-Tibetan 26 pp. 127, 194, 259, 269, 407, 435
Sio 31.3.1
Siona 145 pp. 38, 523
Siouan 104 pp. 115, 548
Siraiya 31.4
Sirenik Yupik, see Yupik (Sirenik) 86
Siriano 145 p. 523
Sirionó 176.6
Siroi 57.14 p. 63
Sisaala 3.5
Sisaala (Western), Ethn. for Sisaala 3.5
Sisiqa 31.3.1
SiSwati, see Swati 3.3.1
Siuslaw 114.3 pp. 210, 407, 490
Siuslawan 114.3
Siwa Berber, see Berber (Siwa) 7.1
Siwa, see Berber (Siwa) 7.1
Siwai, Ethn. for Motuna 81
Siwi, Ethn. for Berber (Siwa) 7.1
Siyin Chin, see Chin (Siyin) 26.2.7
Sko 45 p. 283
Sko 45.3 p. 283
Skou, Ethn. for Sko 45.3
Slave 87.1 pp. 35, 255, 298, 323, 327, 342, 503, 523,
Slavey 87.1
Slavey (North), Ethn., see Slave 87.1
Slavey (Sorth), Ethn., see Slavey 87.1
Slavey (South), Ethn., see Slave 87.1
Slavic 9.10 pp. 58, 63, 190, 223, 291, 523, 573
Slovak 9.10 p. 523
Slovene 9.10 pp. 183, 303
Slovenian, Ethn. for Slovene 9.10
Slovincian 9.10
So 6.3.3 p. 418
Sobei 31.3.1
Sochiapan Chinantec, see Chinantec (Sochiapan) 124.3
Soddo 7.6
Solomons East Papuan 82
Solon 11.2 p. 523
Solor, see Lamaholot 31.2
Somali 7.3.3 pp. 83, 171, 175, 211, 263, 572
Songe 3.3.1
Songhai (Gao), see Koyraboro Senni 6.11
Songhay 6.11
Songhay (Koyra Chiini), Ethn. for Koyra Chiini 6.11
Songhay (Koyraboro Senni), Ethn. for Koyraboro Senni 6.11
Songish 110.2
Soninke 3.10.2 p. 295
Sonsorol-Tobi 31.3.1
Sonsorol, Ethn. for Sonsorol-Tobi 31.3.1
Soo, Ethn. for So 6.3.3
Sooke 110.2
Soqotri 7.6 pp. 35, 83
Sora 29.2
Sorbian 9.10
Sorbian (Lower) 9.10
Sorbian (Lower), Ethn., see Sorbian 9.10
Sorbian (Upper) 9.10
Sorbian (Upper), Ethn., see Sorbian 9.10
Soteapan Zoque, see Zoque (Soteapan) 130
Sotho (Northern) 3.3.1
Sotho (Southern), Ethn. for Sesotho 3.3.1
Sougb 33
Souletin Basque, see Basque (Souletin) 8
South Andamanese 30.2
South Bird's Head 59.2
South Greenlandic, see Greenlandic (South) 86
South Halmahera – West New Guinea 31.3.2
South Korean Sign Language 207 pp. 558, 559, 560, 565
South Mindanao 31.6.7

South Sakhalin Nivkh, see Nivkh (South Sakhalin) 15
South-Central Dravidian 25.3
Southeast Ambrym 31.3.1
Southeastern Pomo, see Pomo (Southeastern) 117.3
Southeastern Tepehuan, see Tepehuan (Southeastern) 108.8
Southern Altai, see Altai (Southern) 11.3
Southern Atlantic 3.2.2
Southern Chukotko-Kamchatkan 14.2
Southern Cushitic 7.3.4
Southern Daly 84.3.5
Southern Dravidian 25.4
Southern Iroquoian 91.2
Southern Ivatan, see Ivatan (Southern) 31.6.4
Southern Khoisan 1.4
Southern Kisi, see Kisi (Southern) 3.2.2
Southern Mam, see Mam (Southern) 132
Southern Paiute, see Paiute (Southern) 108.5
Southern Philippines 31.6.8
Southern Puget Sound Salish, see Salish (Southern Puget Sound) 110.2
Southern Sierra Miwok, see Miwok (Southern Sierra) 113.7.2
Southern Talysh, see Talysh (Southern) 9.8
Southern Tati, see Tati (Southern) 9.8
Southern Teke, see Teke (Southern) 3.3.1
Southern Tiwa, see Tiwa (Southern) 105
Southern Wakashan 109.2
Southern Zulu, see Zulu (Southern) 3.3.1
Southwest Tanna, see Tanna (Southwest) 31.3.1
Southwestern Carriban Creole English, Ethn., see Jamaican Creole 206
Spanish 9.9 pp. 10, 22, 38, 83, 86, 87, 123, 127, 147, 150, 151, 183, 187, 191, 195, 199, 223, 267, 290, 322, 334, 399, 427, 434, 446, 450, 466, 467, 479, 483, 487, 531, 548
Spanish (Canary Islands) 9.9
Spanish Sign Language, Ethn. for Lengua de Señas Española 207
Spanish, Ethn., see Spanish (Canary Islands) 9.9
Spitian 26.2.3
Spokane 110.3 p. 67
Spoken Tamil, see Tamil (Spoken) 25.4
Squamish 110.2 pp. 35, 39, 83, 114, 234, 463
Sranan 206 p. 573
Sre 29.1.2
St. Lawrence Island Yupik, see Yupik (St. Lawrence Island) 86
Standard Spoken Tibetan, see Tibetan (Standard Spoken) 26.2.3
Stieng 29.1.2
Stieng (Budeh), Ethn., see Stieng 29.1.2
Stienh (Bulo), Ethn., see Stieng 29.1.2
Stod Bhoti, Ethn. for Tod 26.2.3
Stoney 104
Stoney, see Dakota 104
Straits Salish, see Salish (Straits) 110.2
Suabo, Ethn. for Inanwatan 59.2
Suau, see Saliba (in Papua New Guinea) 31.3.1
Subiya 3.3.1
Subtiaba-Tlapanec 128
Sudest 31.3.1
Suena 57.4 pp. 123, 343
Sui 28.2 p. 35
Suki 67
Sukuma 3.3.1
Sulawesi 31.6.9
Sulka 77
Sulung 26.2.9
Sumerian p. 569
Sumo Tawahka, Ethn. for Sumu 134
Sumu 134
Sunda, Ethn. for Sundanese 31.6.10
Sundanese 31.6.10
Sundic 31.6.10
Sungor 6.3.9
Supyire 3.5 pp. 42, 171, 230, 295, 323, 343, 447, 531, 548
Sura, see Mupun 7.2.4
Suri, Ethn., see Chai 6.3.8
Suri, Ethn., see Tirmaga 6.3.8
Surmeiran Romansch, see Romansch (Surmeiran) 9.9
Surmic 6.3.8
Sursilvan Romansch, see Romansch (Sursilvan) 9.9
Sursurunga 31.3.1
Susu 3.10.2
Svan 21 p. 318
Svenska Teckenspråket 207 pp. 558, 559
Swahili 3.3.1 pp. 42, 58, 83, 127, 138, 154, 299, 422, 434, 447, 463, 523, 569, 573
Swati 3.3.1
Swedish 9.5 pp. 195, 467
Swedish (Västerbotten) 9.5 p. 247
Swedish Sign Language, Ethn. for Svenska Teckenspråket 207
Swedish, Ethn., see Swedish (Västerbotten) 9.5
Syangja Magar, see Magar (Syangja) 26.2.3
Syrian Arabic, see Arabic (Syrian) 7.6
Taabwa, Ethn. for Tabwa 3.3.1
Taba 31.3.2 pp. 142, 186, 230, 299, 458,
Tabare 57.5
Tabaru 32.3
Tabassaran 20.1.3
Tabla 42
Tabwa 3.3.1
Tacame 156
Tacana 186 pp. 15, 83
Tacanan 186
Tachelhit, Ethn. for Tashlhiyt 7.1
Tagalog 31.6.3 pp. 87, 90, 114, 127, 151, 191, 226, 227, 247, 310, 354, 400, 478
Tagargrent, Ethn. for Berber (Wargla) 7.1
Tagbanwa (Aborlan) 31.6.3
Tahitian 31.3.1 pp. 523, 526
Tai-Kadai 28 pp. 35, 43, 315
Taiap 54
Taiof 31.3.1
Tairora 57.8
Taita, Ethn. for Dabida 3.3.1
Taiwanese 26.1
Taiwanese Sign Language (Ziran Shouyu) 207 pp. 558, 559, 564, 565
Taiwanese Sign Language, Ethn. for Taiwanese Sign Language (Ziran Shouyu) 207
Tajik 9.8
Tajiki, Ethn. for Tajik 9.8
Takelma 115 pp. 210, 318
Takestani, Ethn. for Tati (Southern) 9.8
Takia 31.3.1
Takic 108.6
Talamanca 135.8
Talaud 31.6.9
Talinga 3.3.1 p. 295
Talinga-Bwisi, Ethn. for Talinga 3.3.1
Talodi, Ethn. for Jomang 3.7
Talysh (Azerbaijan) 9.8
Talysh (Southern) 9.8
Talysh, Ethn., see Talysh (Azerbaijan) 9.8
Talysh, Ethn., see Talysh (Southern) 9.8
Tama 6.3.9 p. 83
Tama Sepik 52.4
Tama, Ethn., see Miisiirii 6.3.9
Tamabo 31.3.1
Tamagario 58
Tamahaq (Tahaggart), Ethn. for Tuareg (Ahaggar) 7.1
Tamahaq (Tahaggart), Ethn., see Tuareg (Air) 7.1
Tamahaq (Tahaggart), Ethn., see Tuareg (Ghat) 7.1
Taman 6.3.9
Tamang 26.2.3
Tamang (Eastern), Ethn., see Tamang 26.2.3
Tamazight (Central Atlas), Ethn. for Berber (Middle Atlas) 7.1
Tamazight (Central Atlas), Ethn., see Berber (Ayt Seghrouchen Middle Atlas) 7.1
Tamil 25.4 pp. 127, 130, 131, 162, 174, 231, 258, 442, 451, 483
Tamil (Spoken) 25.4
Tamil, Ethn., see Tamil (Spoken) 25.4
Tampulma 3.5
Tanacross 87.1
Tanaina 87.1 p. 523
Tanana (Lower) 87.1 p. 523
Tanga, see Tangga 31.3.1
Tangale 7.2.4
Tangga 31.3.1
Tangkhul Naga, see Naga (Tangkhul) 26.2.7
Tangkic 84.18 p. 435
Tanglapui 57.19
Tanimbili p. 418
Tanimuca-Retuarã, Ethn. for Retuarã 145
Tanna (Southwest) 31.3.1
Tanzania Sign Language 207 pp. 558, 561, 565
Tanzanian Sign Language, Ethn. for Tanzania Sign Language 207
Taos, see Tiwa (Northern) 105
Tapieté 176.6 p. 523
Tarahumara (Central) 108.7 p. 522
Tarahumara (Western) 108.7
Tarahumara Baja, Ethn. for Tarahumara (Western) 108.7
Tarahumara Central, Ethn. for Tarahumara (Central) 108.7
Tarahumaran 108.7
Tarangan (West) 31.2
Tarao 26.2.7
Taraon, see Digaro 26.2.9
Tarascan 125
Tarascan, see Purépecha 125
Tariana 174 pp. 126, 407
Tariano, Ethn. for Tariana 174
Tarifit, Ethn. for Berber (Rif) 7.1
Tarma Quechua, see Quechua (Tarma) 169
Tarok 3.3.9
Taroko, Ethn. for Seediq 31.1
Tashlhiyt 7.1 p. 47
Tasmanian 85 p. 526
Tasmanian (Oyster Bay to Pitwater) 85
Tat (Muslim) 9.8
Tataltepec Chatino, see Chatino (Tataltepec) 124.8
Tatana, Ethn. for Tatana' 31.6.1
Tatana' 31.6.1
Tatar 11.3 pp. 150, 523
Tatar (Baraba) 11.3
Tatar (Crimean), see Crimean Tatar 11.3
Tatar (Mishar) 11.3
Tatar (Yurt), see Yurt Tatar 11.3
Tatar-Noghay (Alabugat) 11.3
Tatar, Ethn., see Tatar (Baraba) 11.3
Tatar, Ethn., see Tatar (Mishar) 11.3
Tati (Southern) 9.8
Tatuyo 145 p. 523
Taulil 78
Taulil-Butam, Ethn., see Taulil 78
Taushiro 161
Tausug 31.6.3
Tauya 57.14 pp. 155, 175, 198, 235, 263, 350, 351, 362, 422,
Tawala 31.3.1 pp. 223, 406, 414
Tayo 206
Tboli 31.6.7
Teanu, Ethn. for Buma 31.3.1
Teberan 71.2
Teberan-Pawaian 71
Teco 132
Tectiteco, Ethn. for Teco 132
Teda-Daza, see Tubu 6.9
Tedaga, Ethn., see Tubu 6.9
Tegali, Ethn., see Orig 3.7
Tegali, Ethn., see Rashad 3.7
Tehit 32.5
Tehuelche 204.1
Teke (South Central), Ethn. for Ifumu 3.3.1
Teke (Southern) 3.3.1
Telefol 57.18
Telei, see Buin 81
Telugu 25.3 pp. 150, 162
Tem 3.5
Temein 6.3.10
Temiar 29.1.1 pp. 111, 283, 354
Temirgoy Adyghe, see Adyghe (Temirgoy) 19
Temne 3.2.2 p. 83
Ten'a, see Koyukon 87.1
Tenejapa Tzeltal, see Tzeltal (Tenejapa) 132
Tennet 6.3.8 p. 378
Tenyer 3.5
Teop 31.3.1 p. 127
Tepecano 108.8
Tepehua 126 p. 334
Tepehua (Huehuetla), Ethn., see Tepehua 126
Tepehuán (Norte), Ethn. for Tepehuan (Northern) 108.8
Tepehuan (Northern) 108.8

Tepehuan (Southeastern) 108.8 pp. 190, 262, 414
Tepehuán (Sureste), Ethn. for Tepehuan (Southeastern) 108.8
Tepes, see So 6.3.3
Tepetotutla Chinantec, see Chinantec (Tepetotutla) 124.3
Tepiman 108.8
Tequistlatecan 129
Tera 7.2.1
Terêna 174
Teribe 135.8
Ternate 32.3
Teso 6.3.5
Tetela 3.3.1
Tetelcingo Nahuatl, see Nahuatl (Tetelcingo) 108.1
Teton, see Lakhota 104
Tetun 31.2 pp. 142, 503
Tewa (Arizona) 105
Tewa (Rio Grande) 105
Tewa (San Juan Pueblo) 105
Tewa, Ethn. for Tewa (Arizona) 105
Tewa, Ethn., see Tewa (Rio Grande) 105
Tewa, Ethn., see Tewa (San Juan Pueblo) 105
Texmelucan Zapotec, see Zapotec (Texmelucan) 124.8
Tha Pong Nyah Kur, see Nyah Kur (Tha Pong) 29.1.6
Thaayorre 84.17
Thadou 26.2.7
Thai 28.2 pp. 47, 58, 86, 94, 98, 110, 186, 187, 247, 447, 451, 486, 490, 565, 568
Thai Sign Language 207 pp. 558, 559, 561, 565
Thakali 26.2.3
Thami, Ethn. for Thangmi 26.2.3
Thangmi 26.2.3
Thao 31.4
Thargari, see Dhargari 84.17
Thayore, Ethn. for Thaayorre 84.17
Thaypan 84.17
Themne, Ethn. for Temne 3.2.2
Thompson 110.3 pp. 438, 526
Thulung 26.2.3
Thurawal, Ethn. for Dharawal 84.17
Thurgau German, see German (Thurgau) 9.5
Thuringian German, see German (Thuringian) 9.5
Tibetan (Amdo), see Amdo 26.2.3
Tibetan (Modern Literary) 26.2.3
Tibetan (Shigatse) 26.2.3 p. 142
Tibetan (Standard Spoken) 26.2.3 p. 38
Tibetan, Ethn., see Jad 26.2.3
Tibetan, Ethn., see Spitian 26.2.3
Tibetan, Ethn., see Tibetan (Modern Literary) 26.2.3
Tibetan, Ethn., see Tibetan (Shigatse) 26.2.3
Tibetan, Ethn., see Tibetan (Standard Spoken) 26.2.3
Tibeto-Burman 26.2 pp. 43, 95, 235, 283, 287, 299, 315, 355, 359, 363, 394, 569, 573
Ticuna 160
Tiddim Chin, see Chin (Tiddim) 26.2.7
Tidore 32.3 pp. 126, 131
Tifal 57.18
Tigak 31.3.1 pp. 39, 115
Tigemaxo Bozo, see Bozo (Tigemaxo) 3.10.2
Tigré 7.6 pp. 35, 83, 359, 382, 386, 387
Tigré (Beni Amer) 7.6
Tigré, Ethn., see Tigré (Beni Amer) 7.6
Tigrigna, Ethn. for Tigrinya 7.6
Tigrinya 7.6
Tiipay (Jamul) 117.7 pp. 138, 142
Tikar 3.3.1
Tillamook 110.4
Timisoara German, see German (Timisoara) 9.5
Timor-Alor-Pantar 57.19
Timorese 31.2
Timote-Cuica 150
Timote-Cuica, see Cuica 150
Timucua 92
Timugon 31.6.1
Timugon Murut, Ethn. for Timugon 31.6.1
Tinani 26.2.3
Tindi 20.1.1
Tinrin 31.3.1 pp. 262, 414
Tiri, Ethn. for Tinrin 31.3.1
Tiriyo 175
Tirmaga 6.3.8
Tiruray 31.6.7
Tiv 3.3.1
Tiwa (Northern) 105
Tiwa (Southern) 105
Tiwi 84.19 pp. 78, 102, 142, 226, 242, 475, 478
Tiwian 84.19
Tlahuitoltepec Mixe, see Mixe (Tlahuitoltepec) 130
Tlapanec 128
Tlapaneco (Acatepec), Ethn., see Tlapanec 128
Tlingit 87.3 pp. 14, 26, 30, 34, 38, 39, 78, 171, 407
To'abaita, Ethn. for Toqabaqita 31.3.1
Toaripi 72
Toba 199
Toba Batak, see Batak (Toba) 31.6.10
Tobati 31.3.1 p. 331
Tobelo 32.3
Tod 26.2.3
Toda 25.4
Tofa 11.3
Tofalar, see Tofa 11.3
Togo Gurma, see Gurma (Togo) 3.5
Tojolabal 132
Tok Pisin 206
Tokelauan 31.3.1
Tol 133
Tolai 31.3.1 pp. Oceanic; New Britain, PNG
Tondano 31.6.9 p. 474
Tonga (in Zambia) 3.3.1 p. 210
Tonga, Ethn. for Tonga (in Zambia) 3.3.1
Tongan 31.3.1 pp. 170, 222, 226
Tonkawa 99 p. 338
Tontemboan 31.6.9
Toqabaqita 31.3.1 p. 226
Tor 40
Toraja 31.6.9
Toraja-Sa'dan, Ethn. for Toraja 31.6.9
Toratán 31.6.9
Toro So Dogon, see Dogon (Toro So) 3.4
Torres Strait Creole, Ethn. for Broken 206
Torricelli 51 p. 111
Torwali 9.7
Totonac (Misantla) 126 pp. 138, 414, 415
Totonac (Papantla) 126
Totonac (Sierra) 126
Totonac (Xicotepec de Juárez) 126
Totonaca (Papantla), Ethn. for Totonac (Papantla) 126
Totonaca (Xicotepec de Juarez), Ethn. for Totonac (Xicotepec de Juárez) 126
Totonaca (Yecuatla), Ethn. for Totonac (Misantla) 126
Totonacan 126
Totontepec Mixe, see Mixe (Totontepec) 130
Toussian 3.5
Toussian (Northern), Ethn., see Toussian 3.5
Toussian (Southern), Ethn. for Toussian (Win) 3.5
Toussian (Southern), Ethn., see Toussian 3.5
Toussian (Win) 3.5
Towa, see Jemez 105
Trans-New Guinea 57 pp. 122, 194, 327
Trió, Ethn. for Tiriyo 175
Tripuri, see Kokborok 26.2.2
Trique (Chicahuaxtla) 124.4
Trique (Copala) 124.4
Trukese, see Chuukese 31.3.1
Trumai 178 pp. 183, 202, 203, 263, 287, 407, 419, 514
Trumaí, Ethn. for Trumai 178
Trung, see Dulong 26.2.11
Tsafiki 140
Tsakhur 20.1.3
Tsamosan 110.5
Tsat 31.6.10
Tschama, see Shipibo-Konibo 168
Tsez 20.1.1 pp. 127, 135, 438
Tshangla 26.2.3
Tsimané, Ethn. for Mosetén 190
Tsimshian (Coast) 113.6 pp. 30, 222, 226, 230, 394, 395
Tsimshian, Ethn., see Tsimshian (Coast) 113.6
Tsimshianic 113.6 p. 11
Tsogo 3.3.1
Tsonga 3.3.1
Tsou 31.5 p. 23
Tsouic 31.5
Tsova-Tush 20.2 pp. 210, 315, 318
Tswana 3.3.1 p. 523
Tu, Ethn. for Mangghuer 11.1
Tuamotuan 31.3.1 pp. 523, 526
Tuareg (Ahaggar) 7.1 pp. 490, 569
Tuareg (Air) 7.1
Tuareg (Ghat) 7.1
Tuareg (Tahaggart), see Tuareg (Ahaggar) 7.1
Tubar 108.9 p. 522
Tubarão, Ethn. for Aikaná 174
Tübatulabal 108.10
Tubu 6.9 pp. 258, 486, 490
Tucano 145 p. 523
Tucanoan 145 pp. 319, 327, 523
Tugun 31.2 p. 503
Tukang Besi 31.6.9 pp. 234, 235, 255, 283, 379, 391, 398, 400, 434, 442, 443
Tukangbesi (North), Ethn., see Tukang Besi 31.6.9
Tukangbesi (South), Ethn., see Tukang Besi 31.6.9
Tuki 3.3.1
Tukpa, Ethn. for Nyamkad 26.2.3
Tulu 25.4
Tumak 7.2.2
Tumleo 31.3.1
Tümpisa Shoshone 108.5 pp. 107, 147, 174, 175, 346, 522, 523
Tumucua, see Timucua 92
Tumzabt, Ethn. for Berber (Mzab) 7.1
Tundra Nenets, see Nenets (Tundra) 10.2
Tundra Yukaghir, see Yukaghir (Tundra) 12
Tunebo 135.2
Tunebo (Angosturas), Ethn., see Tunebo 135.2
Tunen 3.3.1 p. 343
Tungag, Ethn. for Tungak 31.3.1
Tungak 31.3.1
Tungusic 11.2 pp. 42, 548, 549
Tunica 95 pp. 451, 470, 526
Tunisian Arabic, see Arabic (Tunisian) 7.6
Tupari 176.5
Tupi 176.6 p. 523
Tupi-Guaraní 176.6
Tupian 176 p. 523
Tupuri 3.1
Turama-Kikorian 70
Turinese Italian, see Italian (Turinese) 9.9
Türk Isaret Dili 207 p. 558
Turkana 6.3.5 pp. 210, 282, 323, 548
Turkic 11.3 pp. 42, 95, 147, 291, 298, 299, 314, 319, 415, 487, 548, 549
Turkic (East-Central Xorasan) 11.3
Turkic (West Xorasan) 11.3
Turkish 11.3 pp. 15, 43, 50, 86, 87, 90, 95, 115, 146, 147, 151, 158, 174, 194, 206, 219, 290, 311, 314, 315, 318, 334, 338, 407, 434, 446, 487, 495, 496, 502, 523, 531, 546, 569
Turkish Sign Language, Ethn. for Türk Isaret Dili 207
Turkmen 11.3 p. 523
Tuscarora 91.1
Tutchone (Northern) 87.1 p. 523
Tutelo 104
Tuvaluan 31.3.1 pp. 223, 450
Tuvaluan, Ethn., see Nanumea 31.3.1
Tuvan 11.3 pp. 114, 307, 548
Tuvin, Ethn. for Tuvan 11.3
Tuyuca 145 pp. 314, 523
Twana 110.2
Twi, see Akan 3.9
Tyap, Ethn., see Gworok 3.3.9
Tyaraity, Ethn. for Tyeraity 84.3.4
Tyeraity 84.3.4
Tyvan, see Tuvan 11.3
Tzeltal 132
Tzeltal (Aguacatenango) 132
Tzeltal (Bachajón) 132
Tzeltal (Bachajón), Ethn., see Tzeltal 132
Tzeltal (Oxchuc), Ethn., see Tzeltal 132
Tzeltal (Oxchuc), Ethn., see Tzeltal (Aguacatenango) 132
Tzeltal (Oxchuc), Ethn., see Tzeltal (Tenejapa) 132
Tzeltal (Tenejapa) 132
Tzotzil 132
Tzotzil (Chamula), Ethn., see Tzotzil 132
Tzotzil (San Andrés Larrainzar), Ethn. for Tzotzil (San Andrés) 132
Tzotzil (San Andrés) 132 p. 526
Tzotzil (Zinacantán) 132
Tzutujil 132 pp. 98, 106, 239, 438, 467, 510

Tzutujil (Eastern), Ethn., see Tzutujil 132
Tzutujil (Western), Ethn., see Tzutujil 132
Ubde-Nehern, see Hupda 173
Ubykh 19 pp. 111, 258
Udi 20.1.3
Udihe 11.2 pp. 143, 223, 394, 523
Udmurt 10.1.1
Ufromi, see Gorowa 7.3.4
Ugandan Sign Language 207 pp. 558, 559, 560, 561
Ugandian Sign Language, Ethn. for Ugandan Sign Language 207
Ugric 10.1.2
Uigur, see Uyghur 11.3
Ukrainian 9.10 p. 523
Ulch, Ethn. for Ulcha 11.2
Ulcha 11.2 p. 42
Uldeme 7.2.1
Ulithian 31.3.1
Ulu Muar Malay, see Malay (Ulu Muar) 31.6.10
Uma 31.6.9
Uma' Lung Kenyah, see Kenyah (Uma' Lung) 31.6.1
Umatilla Sahaptin, see Sahaptin (Umatilla) 113.5
Umatilla, Ethn. for Sahaptin (Umatilla) 113.5
Umaua 175
Umbu-Ungu, Ethn. for Kaugel 57.5
UMbundu 3.3.1 p. 23
Umpqua, see Siuslaw 114.3
Una 57.17 pp. 27, 147
Unami 89.1
Uneapa, Ethn., see Bali-Vitu 31.3.1
Ungarinjin 84.21 pp. 183, 211, 235, 334, 451
Upper Austrian German, see German (Upper Austrian) 9.5
Upper Chehalis, see Chehalis (Upper) 110.5
Upper Chinook, see Chinook (Upper) 113.1
Upper Kuskokwim 87.1
Upper Sepik 52.5
Upper Sorbian, see Sorbian (Upper) 9.10
Upper Yuat 56
Ura 31.3.1
Uradhi 84.17 p. 98
Urak Lawoi' 31.6.10
Uralic 10 pp. 50, 91, 94, 127, 147, 307, 323, 415, 487, 548
Urarina 164
Urdu 9.7 p. 569
Urhobo 3.3.4
Urim 51.2
Uripiv-Wala-Rano-Atchin, Ethn. for Atchin 31.3.1
Uru 193
Uru-Chipaya 193
Urubú Sign Language 207 pp. 558, 565
Urubú-Kaapor 176.6 pp. 191, 523, 526
Urubú-Kaapor Sign Language, Ethn. for Urubú Sign Language 207
Urum 11.3
Usan 57.14 pp. 38, 67, 147, 415, 548
Usarufa 57.8
Ute 108.5 pp. 154, 434, 522, 523
Ute-Southern Paiute, Ethn., see Chemehuevi 108.5
Ute-Southern Paiute, Ethn., see Paiute (Southern) 108.5
Ute-Southern Paiute, Ethn., see Ute 108.5
Utian 113.7 pp. 548, 549
Uto-Aztecan 108 pp. 107, 115, 194, 199, 315, 323, 414, 442, 491, 522, 523, 548
Uyghur 11.3 p. 568
Uzbek 11.3 pp. 147, 490, 523
Uzbek (Northern) 11.3
Uzbek (Northern), Ethn., see Krymchak 11.3
Uzbek (Northern), Ethn., see Uzbek 11.3
Uzbek (Southern), Ethn., see Uzbek 11.3
Vafsi 9.8
Vagla 3.5
Vai 3.10.2 pp. 110, 295, 447, 487, 567
Valman, Ethn. for Walman 51.3
Vanimo, Ethn. for Dumo 45.3
Vasavi 9.7
Västerbotten Swedish, see Swedish (Västerbotten) 9.5
Vata 3.8
Vaupés-Japurá 173
Vedda 9.7 p. 523
Veddah, Ethn. for Vedda 9.7
Venda 3.3.1 p. 178
Vengo, Ethn. for Babungo 3.3.1
Veps 10.1.1
Viennese German, see German (Viennese) 9.5
Viet-Muong 29.1.9
Vietnamese 29.1.9 pp. 26, 35, 43, 58, 94, 95, 127, 186, 187, 202, 218, 222, 226, 227, 250, 290, 447, 451
Vlaams, Ethn., see Dutch (Zeeuws) 9.5
Vlaamse Gebarentaal 207 pp. 558, 559
Vod, Ethn. for Votic 10.1.1
Vogul, see Mansi 10.1.2
Votic 10.1.1
Wa, see Parauk 29.1.8
Waalubal Bandjalang, see Bandjalang (Waalubal) 84.17
Waata Oromo, see Oromo (Waata) 7.3.3
Waata, see Oromo (Waata) 7.3.3
Wadjiginy, Ethn. for Bachamal 84.3.1
Wageman, Ethn. for Wagiman 84.7.8
Wagiman 84.7.8
Wahehe, see Hehe 3.3.1
Wahgi 57.5 p. 79
Wahgi (North), Ethn., see Wahgi 57.5
Wahgi, Ethn., see Wahgi 57.5
Wai Wai 175 p. 223
Waigali 9.7 p. 548
Waimaha, Ethn. for Barasano (Northern) 145
Waiwai, Ethn. for Wai Wai 175
Wajarri, Ethn. for Watjarri 84.17
Wakami, see Kami 3.3.1
Wakashan 109 pp. 11, 35, 175
Wakhi 9.8
Walapai, see Hualapai 117.7
Walla Walla, Ethn., see Sahaptin (Northern) 113.5
Wallisian 31.3.1
Walmajarri, Ethn. for Walmatjari 84.17
Walman 51.3
Walmatjari 84.17
Waluwara, see Warluwara 84.17
Wambaya 84.20 pp. 118, 443, 526
Wambon 57.3 p. 139
Wamei, Ethn. for Konyagi 3.2.1
Wampanoag, Ethn. for Massachusett 89.1
Wandala, Ethn., see Malgwa 7.2.1
Wandamen p. 183
Wandarang, Ethn. for Warndarang 84.13
Wangaaybuwan-Ngiyambaa, Ethn., see Ngiyambaa 84.17
Wangkumara 84.17
Wantoat 57.10
Waorani 157 p. 534
Wapei-Palei 51.3
Wapishana 174
Wapogoro, see Pogoro 3.3.1
Wappo 120.1
Wappo-Yukian 120
Warao 153 pp. 46, 331
Waray 84.7.7
Waray (in Australia) 84.7.7
Waray-Waray 31.6.3
Waray, Ethn. for Waray (in Australia) 84.7.7
Wardaman 84.7.8 pp. 131, 147, 262, 307, 347, 438
Warekena 174 pp. 182, 226, 291, 470
Warembori 36
Wargamay, see Warrgamay 84.17
Wargla Berber, see Berber (Wargla) 7.1
Wargla, see Berber (Wargla) 7.1
Wari' 182 pp. 46, 50, 146, 175, 363, 370, 399, 400, 431, 548
Waris 46
Warlpiri 84.17 p. 526
Warluwara 84.17
Warndarang 84.13 pp. 242, 463, 495, 496
Waropen 31.3.2
Warrgamay 84.17
Warrnambool 84.17
Warrungu 84.17 pp. 398, 438
Warrwa 84.16
Warungu, Ethn. for Warrungu 84.17
Wasco-Wishram, see Chinook (Upper) 113.1
Washkuk, see Kwoma 52.1
Washo 119 pp. 43, 239
Wasi, Ethn. for Alagwa 7.3.4
Waskia 57.14 p. 482
Watam 53.3
Wathawurrung 84.17
Watjarri 84.17
Waunana 136 p. 294
Waurá 174
Wayampi 176.6 p. 523
Wayampi (Amapari), Ethn. for Wayampi 176.6
Wayampi (Oiapoque), Ethn. for Wayampi 176.6
Wayan Fijian, see Fijian (Wayan) 31.3.1
Wayana 175
Wayãpi, see Wayampi 176.6
Wayu, Ethn. for Hayu 26.2.3
Wayuu, Ethn. for Goajiro 174
Wè Northern, Ethn. for Wobe 3.8
Wè Southern, Ethn., see Guere 3.8
Wè Western, Ethn., see Guere 3.8
Wea, see Illinois 89.1
Wedau 31.3.1
Welsh 9.4 pp. 199, 211, 307, 334, 363, 415
Welsh Romani, see Romani (Welsh) 9.7
Wembawemba 84.17
Weme-Gbe, Ethn. for Wéménugbé 3.9
Wéménugbé 3.9
Weri 57.11 p. 62
Wersing, Ethn. for Kolana 57.19
West Barkly 84.20
West Bird's Head 32.5
West Bomberai 34
West Bougainville 80
West Chadic 7.2.4
West Greenlandic, see Greenlandic (West) 86
West Makian 32.3
West Papuan 32 pp. 111, 183
West Tarangan, see Tarangan (West) 31.2
West Wapei 51.4
West Xorasan Turkic, see Turkic (West Xorasan) 11.3
West Yugur, Ethn. for Saryg Yughur 11.3
West-Central Oromo, see Oromo (West-Central) 7.3.3
Western Abenaki, see Abenaki (Western) 89.1
Western Apache, see Apache (Western) 87.1
Western Armenian, see Armenian (Western) 9.2
Western Arrernte, see Arrernte (Western) 84.17
Western Bru, see Bru (Western) 29.1.3
Western Bukidnon Manobo, see Manobo (Western Bukidnon) 31.6.8
Western Cham, see Cham (Western) 31.6.10
Western Daly 84.3.6
Western Dangaléat, see Dangaléat (Western) 7.2.2
Western Desert (Ooldea) 84.17
Western Desert, see Ngaanyatjarra 84.17
Western Fly 69
Western Frisian, see Frisian (Western) 9.5
Western Kanjobal, see Kanjobal (Western) 132
Western Malayo-Polynesian 31.6
Western Mande 3.10.2
Western Maninka, see Maninka (Western) 3.10.2
Western Sko 45.3
Western Tarahumara, see Tarahumara (Western) 108.7
Westphalian German, see German (Westphalian) 9.5
Westphalien, Ethn. for German (Westphalian) 9.5
Wetan 31.2
Wichí 198 pp. 546, 548
Wichí Lhamtés Güisnay, Ethn., see Wichí 198
Wichita 103 pp. 78, 94, 179, 399, 400, 407
Wik Munkan 84.17
Wik Ngathana 84.17 p. 331
Wik-Mungkan, Ethn. for Wik Munkan 84.17
Wik-Ngathana, Ethn. for Wik Ngathana 84.17
Wikchamni 113.9 p. 43
Win Toussian, see Toussian (Win) 3.5
Wind River Shoshone, see Shoshone (Wind River) 108.5
Windesi p. 183
Winnebago 104 pp. 62, 63, 526
Wintu 113.8 pp. 34, 39, 526, 546, 548
Wintu, Ethn., see Patwin 113.8
Wintuan 113.8 pp. 548, 549
Winye Ko, see Ko (Winye) 3.5
Winyé, Ethn. for Ko (Winye) 3.5
Wirangu 84.17
Wishram, see Chinook (Upper) 113.1
Wissel Lakes-Kemandoga 57.20
Wiyot 89.2
Wo'tiheh, see Piaroa 143
Wobe 3.8

Wogamusin 52.5
Woisika 57.19
Woiwurrung 84.17
Wolaytta 7.5 p. 415
Woleaian 31.3.1 p. 414
Wolio 31.6.9
Wolof 3.2.1 pp. 50, 58, 235
Womo 45.1
Worora 84.21 p. 183
Wororan 84.21
Woun Meu, Ethn. for Waunana 136
Wu (Changzhou) 26.1
Wuzlam, Ethn. for Uldeme 7.2.1
Wyandot 91.1
Xakas, see Khakas 11.3
Xalitla Nahuatl, see Nahuatl (Xalitla) 108.1
/Xam 1.4
Xârâcùù 31.3.1
Xasonga 3.10.2
Xavánte 177.2
Xerénte 177.2
Xhosa 3.3.1 p. 573
Xiamen 26.1 p. 38
Xicotepec de Juárez Totonac, see Totonac (Xicotepec de Juárez) 126
Xokleng 177.2 p. 526
!Xóõ 1.4 pp. 10, 35, 82, 374, 522, 554
!Xoo, Ethn. for !Xóõ 1.4
Xorasan Turkic, see Turkic (East-Central Xorasan) 11.3
Xorasan, see Turkic (West Xorasan) 11.3
!Xu, see Ju|'hoan 1.2
!Xun (Ekoka) 1.2 p. 327
Yabem, Ethn. for Jabêm 31.3.1
Yag Dii 3.1
Yagan, see Yahgan 205
Yagaria 57.8 pp. 523, 526
Yagaria, Ethn., see Hua 57.8
Yaghnobi 9.8
Yagnobi, Ethn. for Yaghnobi 9.8
Yagua 159 pp. 94, 98, 163, 165, 250, 269, 415
Yahgan 205
Yaitepec Chatino, see Chatino (Yaitepec) 124.8
Yaka 3.3.1 p. 523
Yaka (in Central African Republic), Ethn. for Aka 3.3.1
Yaka (in Democratic Republic of Congo), Ethn. for Yaka 3.3.1
Yakan 31.6.6
Yakima, Ethn., see Sahaptin (Northern) 113.5
Yakoma 3.1 pp. 290, 291
Yakut 11.3 p. 50
Yale 49
Yale (Kosarek) 57.17
Yale, Ethn. for Nagatman 49
Yali 57.6
Yali (Angguruk), Ethn., see Yali 57.6
Yámana 205
Yámana, Ethn. for Yahgan 205
Yamba 3.3.1
Yamdena 31.2
Yami 31.6.4 p. 222
Yaminahua 168
Yana 117.6 p. 526
Yanesha', Ethn. for Amuesha 174
Yangmanic 84.7.8
Yankuntjatjara 84.17
Yankunytjatjara, Ethn. for Yankuntjatjara 84.17
Yanomam 154
Yanomámi 154
Yansi 3.3.1
Yanyula, see Yanyuwa 84.17
Yanyuwa 84.17
Yao (in Malawi) 3.3.1
Yao, Ethn. for Yao (in Malawi) 3.3.1
Yao, see Punu 27
Yapese 31.6.11 pp. 34, 35, 66, 83, 418, 526
Yaqay 59.1
Yaqui 108.2 pp. 91, 390, 548
Yareba 73 p. 54
Yareban 73
Yaruro 152
Yaruru, Ethn. for Yaruro 152
Yatê 177.8
Yatê, see Fulniô 177.8
Yatzachi Zapotec, see Zapotec (Yatzachi) 124.8
Yaudanchi Yokuts, see Yokuts (Yaudanchi) 113.9
Yavapai 117.7
Yawa 35 p. 414
Yawelmani 113.9 pp. 115, 548
Yawuru 84.16 p. 211
Yay 28.2 pp. 83, 451
Yaygir 84.17
Yazgulyam 9.8
Yazva 10.1.1
Yebamasa, see Barasano 145
Yei 63
Yekhee, Ethn. for Etsako 3.3.4
Yele 75
Yele, Ethn. for Yelî Dnye 75
Yeletnye, see Yelî Dnye 75
Yelî Dnye 75 pp. 79, 83
Yellow River 52.6
Yelmek 62
Yemba 3.3.1
Yemsa 7.5
Yeniseian 13
Yessan-Mayo 52.4 p. 183
Yeyi 3.3.1 pp. 11, 82
Yi 26.2.4 p. 569
Yi (Central), Ethn., see Yi 26.2.4
Yi (Western), Ethn. for Lalo 26.2.4
Yiddish 9.5
Yiddish (Bessarabian) 9.5
Yiddish (Eastern), Ethn., see Yiddish 9.5
Yiddish (Eastern), Ethn., see Yiddish (Bessarabian) 9.5
Yiddish (Eastern), Ethn., see Yiddish (Lithuanian) 9.5
Yiddish (Eastern), Ethn., see Yiddish (Lodz) 9.5
Yiddish (Lithuanian) 9.5
Yiddish (Lodz) 9.5 p. 223
Yiddish (Western), Ethn., see Yiddish 9.5
Yidiny 84.17 pp. 230, 342, 347, 438, 531
Yil 51.3
Yimas 53.4 pp. 14, 43, 79, 206, 207, 303, 422, 423
Yindjibarndi 84.17 p. 71
Yine, Ethn. for Piro 174
Yinggarda, Ethn. for Yingkarta 84.17
Yingkarta 84.17
Yir Yiront 84.17 p. 526
Yokuts 113.9
Yokuts (Valley), Ethn., see Yawelmani 113.9
Yokuts (Yaudanchi) 113.9
Yokuts, Ethn., see Wikchamni 113.9
Yokuts, Ethn., see Yokuts (Yaudanchi) 113.9
Yolngu, see Djapu 84.17
Yongkom, Ethn., see Kati (in West Papua, Indonesia) 57.18
Yorta Yorta 84.17
Yoruba 3.3.3 pp. 22, 46, 58, 83, 102, 127, 143, 226, 255, 271, 447, 554
Yosondúa Mixtec, see Mixtec (Yosondúa) 124.4
Yucatec 132 p. 35
Yuchi 93 p. 39
Yucuna 174
Yugambeh, see Bandjalang (Yugumbir) 84.17
Yugh 13
Yughur (Saryg), see Saryg Yughur 11.3
Yugumbir Bandjalang, see Bandjalang (Yugumbir) 84.17
Yukaghir 12 p. 479
Yukaghir (Kolyma) 12 pp. 42, 146, 206, 250, 367, 382, 390, 430, 431, 454, 475, 514, 548
Yukaghir (Northern), Ethn. for Yukaghir (Tundra) 12
Yukaghir (Southern), Ethn. for Yukaghir (Kolyma) 12
Yukaghir (Tundra) 12
Yuki 120.2
Yukian 120.2
Yukpa 175 pp. 523, 526
Yukulta 84.18 pp. 178, 338, 382, 406, 407, 422, 438, 439, 526
Yulparija 84.17 pp. 422, 522
Yulu 6.2.1 pp. 35, 83
Yulu, Ethn., see Binga 6.2.1
Yuma, see Quechan 117.7
Yuman 117.7 pp. 115, 400, 548
Yup'ik (Central) 86 pp. 43, 110, 118, 311, 438
Yup'ik (Central), Ethn., see Yup'ik (Norton Sound) 86
Yup'ik (Chevak) 86
Yup'ik (Norton Sound) 86
Yupik (Central Siberian), Ethn., see Yupik (Siberian) 86
Yupik (Central Siberian), Ethn., see Yupik (St. Lawrence Island) 86
Yupik (Central), Ethn. for Yup'ik (Central) 86
Yupik (Central), Ethn., see Yup'ik (Chevak) 86
Yupik (Naukan) 86
Yupik (Pacific Gulf), Ethn. for Alutiiq 86
Yupik (Siberian) 86 p. 22
Yupik (Sirenik) 86
Yupik (St. Lawrence Island) 86 p. 67
Yuracare 191
Yuri, see Karkar-Yuri 48
Yurimangí 138
Yurok 89.3 pp. 34, 35, 118, 147, 311, 374
Yurt Tatar 11.3
Yuruti 145 p. 523
Yuulngu, see Gaalpu 84.17
Yuwaalaraay 84.17 p. 74
Zabwa, see Tabwa 3.3.1
Zamucoan 196
Zande 3.1
Zaparo 158
Záparo, Ethn. for Zaparo 158
Zaparoan 158
Zapateco (Ixtlán Sureste), Ethn. for Zapotec (Ixtlan) 124.8
Zapotec (Isthmus) 124.8
Zapotec (Ixtlan) 124.8
Zapotec (Juárez) 124.8
Zapotec (Mitla) 124.8
Zapotec (Mixtepec) 124.8
Zapotec (Quiegolani) 124.8
Zapotec (San Juan Guelavía), Ethn., see Zapotec (San Lucas Quiaviní) 124.8
Zapotec (San Lucas Quiaviní) 124.8
Zapotec (Texmelucan) 124.8
Zapotec (Yatzachi) 124.8
Zapotec (Zoogocho) 124.8
Zapotecan 124.8
Zapoteco (Istmo), Ethn. for Zapotec (Isthmus) 124.8
Zapoteco (Juárez Sierra), Ethn. for Zapotec (Juárez) 124.8
Zapoteco (Mitla), Ethn. for Zapotec (Mitla) 124.8
Zapoteco (Mixtepec), Ethn. for Zapotec (Mixtepec) 124.8
Zapoteco (Santa María Quiegolani), Ethn. for Zapotec (Quiegolani) 124.8
Zapoteco (Texmelucan), Ethn. for Zapotec (Texmelucan) 124.8
Zarma 6.11
Zayse 7.5 p. 470
Zayse-Zergulla, Ethn. for Zayse 7.5
Zazaki 9.8
Zeberio Basque, see Basque (Zeberio) 8
Zeeuws Dutch, see Dutch (Zeeuws) 9.5
Zeme Naga, see Naga (Zeme) 26.2.7
Zenaga 7.1
Zhang-Zhung 26.2.3
Zhanglu Kam, see Kam (Zhanglu) 28.2
Zhuang (Northern) 28.2
Zimakani 59.1
Zinacantán Tzotzil, see Tzotzil (Zinacantán) 132
Ziran Shouyu Taiwanese Sign Language, see Taiwanese Sign Language (Ziran Shouyu) 207
Zoogocho Zapotec, see Zapotec (Zoogocho) 124.8
Zoque (Chimalapa) 130
Zoque (Copainalá) 130 pp. 438, 463
Zoque (Copainalá), Ethn., see Zoque (Ostuacan) 130
Zoque (Francisco León) 130
Zoque (Ostuacan) 130 p. 334
Zoque (Rayon) 130
Zoque (Soteapan) 130
Zulu 3.3.1 pp. 11, 35, 82, 151, 211, 290, 502, 523
Zulu (Northern) 3.3.1
Zulu (Southern) 3.3.1
Zulu, Ethn., see Zulu (Northern) 3.3.1
Zulu, Ethn., see Zulu (Southern) 3.3.1
Zuni 107 pp. 22, 434
Zurich German, see German (Zurich) 9.5
Zyrian, see Komi-Zyrian 10.1.1

WALS Code Index

Listed here are the three-letter codes used on the maps, followed by the *WALS* name of the language. See §3.5 of the Introduction for a description of some of the principles used for assigning these codes. Each language name is followed by its genus number (see the Genealogical Language List).

aab Arapesh (Abu) 51.1
aar Aari 7.5
aba Abau 52.5
abb Arabic (Abbéché Chad) 7.6
abd Abidji 3.9
abe Arabic (Beirut) 7.6
abh Arabic (Bahrain) 7.6
abi Abipón 199
abk Abkhaz 19
abm Alabama 94
abn Arabana 84.17
abo Arbore 7.3.3
abu Abun 32.4
abw Abenaki (Western) 89.1
abz Abaza 19
ace Acehnese 31.6.10
acg Achagua 174
ach Aché 176.6
aci Achí 132
acl Acholi 6.3.5
acm Achumawi 117.2
acn Achang 26.2.4
aco Acoma 106
acu Achuar 162
ada Adamorobe Sign Language 207
adi Adioukrou 3.9
adk Andoke 147
adn Adynyamathanha 84.17
adt Adyghe (Temirgoy) 19
ady Adyghe (Abzakh) 19
adz Adzera 31.3.1
aea Aleut (Eastern) 86
aeg Arabic (Egyptian) 7.6
ael Arabic (Eastern Libyan) 7.6
afr Afrikaans 9.5
aga Agarabi 57.8
agc Agta (Central) 31.6.4
agh Aghem 3.3.1
agl Aghul 20.1.3
agm Angami 26.2.7
ago Angolar 206
agr Aguaruna 162
agt Anguthimri 84.17
agu Aguacatec 132
agw Alagwa 7.3.4
aho Arapaho 89.1
ahs Arabic (Bani-Hassan) 7.6
aht Ahtna 87.1
ahu Aghu 57.3
aik Aikaná 174
ain Ainu 16
aiz Aizi 3.8
aja Aja 6.2.2
aji Ajië 31.3.1
aka Aka 3.3.1
akb Aka-Biada 30.1
akc Aka-Cari 30.1
akh Akha 26.2.4
akk Aka-Kede 30.1
akl Aklanon 31.6.3
akm Arakanese (Marma) 26.2.4
akn Akan 3.9
ako Arabic (Kormakiti) 7.6
akr Arikara 103
akw Akawaio 175
ala Alamblak 52.3
alb Albanian 9.1
alc Allentiac 201
ald Alladian 3.9
ale Aleut 86
alg Algonquin 89.1
alk Arop-Lokep 31.3.1
all Ala'ala 31.3.1
aln Alune 31.2
als Alsea 114.1
alt Alsatian 9.5
alu Alutor 14.1
alw Alawa 84.13
alx Alas 31.6.10
aly Alyawarra 84.17
ama Amara 31.3.1
amb Ambulas 52.1
amc Amahuaca 168
amd Amdo 26.2.3
ame Amele 57.14
amh Amharic 7.6
ami Amis 31.4
amk Amarakaeri 171
aml Ambae (Lolovoli Northeast) 31.3.1
amn Amanab 46
amo Amo 3.3.7
amp Arrernte (Mparntwe) 84.17
amq Ambai 31.3.2
amr Arabic (Moroccan) 7.6
ams Arabic (Modern Standard) 7.6
amu Amuesha 174
amx Anamuxra 57.14
amz Amuzgo 124.1
ana Araona 186
anc Angas 7.2.4
ane Anem 76
ang Anggor 47
ani //Ani 1.1
anj Anejom 31.3.1
anl Arabic (North Levantine Spoken) 7.6
ann Anindilyakwa 84.7.1
ano Anong 26.2.11
ant Angaatiha 57.1
anu Anufo 3.9
anx Andi 20.1.1
any Anywa 6.3.5
ao Ao 26.2.7
apa Arabic (Palestinian) 7.6
apc Apache (Chiricahua) 87.1
ape Ampeeli 57.1
api Apinayé 177.2
apj Apache (Jicarilla) 87.1
apk A-Pucikwar 30.1
apl Apalaí 175
apt Apatani 26.2.9
apu Apurinã 174
apw Apache (Western) 87.1
ara Arawak 174
arb Arabela 158
arc Archi 20.1.3
ard Arandai 59.2
arg Arabic (Gulf) 7.6
arh Arabic (Hijazi) 7.6
ari Aribwatsa 31.3.1
ark Araki 31.3.1
arl Arabic (Lebanese) 7.6
arm Armenian (Eastern) 9.2
arn Arabic (Borno Nigerian) 7.6
aro Arosi 31.3.1
arp Arapesh 51.1
arq Arabic (Iraqi) 7.6
arr Arrernte 84.17
ars Arabic (San'ani) 7.6
arv Arabic (Negev) 7.6
arw Armenian (Western) 9.2
arx 'Are'are 31.3.1
arz Armenian (Iranian) 9.2
ash Adyghe (Shapsugh) 19
asl American Sign Language 207
asm Asmat 57.2
aso Altai (Southern) 11.3
ass Assamese 9.7
ast Asturian 9.9
asu Asuriní 176.6
asy Arabic (Syrian) 7.6
ata Atayal 31.1
atb Aralle-Tabulahan 31.6.9
atc Atchin 31.3.1
ath Athpare 26.2.3
ati Atikamekw 89.1
atk Atakapa 97
atm Atacameño 194
atq Alutiiq 86
ats Atsugewi 117.2
atu Arabic (Tunisian) 7.6
au Au 51.3
aul Aulua 31.3.1
aus Auslan 207
auy Auyana 57.8
ava Avar 20.1.1
avo Avokaya 6.2.6
avt Avatime 3.9
awa Awa 57.8
awd Awadhi 9.7
awe Arrernte (Western) 84.17
awi Awin 64
awk Akwa 3.3.1
awn Awngi 7.3.2
awp Awa Pit 140
awt Awtuw 52.2
awy Awyi 46
axv Akhvakh 20.1.1
ayi Anyi 3.9
aym Aymara 170
ayn Aynu 11.3
ayo Ayomán 149
ayr Ayoreo 196
ayw Ayiwo 83
aze Azerbaijani 11.3
azi Azari (Iranian) 11.3
baa Barai 57.12
bab Babungo 3.3.1
bac Bachamal 84.3.1
bad Bade 7.2.4
bae Baré 174
baf Bafut 3.3.1
bag Bagirmi 6.2.1
bah Bahuana 174
bai Bai 26.2.1
baj Bajau 31.6.6
bak Baka (in Cameroon) 3.1
bal Balinese 31.6.10
bam Bambara 3.10.2
ban Bana 7.2.1
bao Baonan 11.1
baq Baure 174
bar Bari 6.3.5
bas Basaá 3.3.1
bat Batak 31.6.3
bau Bau 57.14
baw Bawm 26.2.7
bbf Bobo Fing 3.10.2
bbl Babole 3.3.1
bbm Bambam 31.6.9
bbu Barambu 3.1
bbw Bininj Gun-Wok 84.7.2
bca Bandjalang (Casino) 84.17
bch Berber (Chaouia) 7.1
bco Bella Coola 110.1
bdc Berbice Dutch Creole 206
bdg Badaga 25.4
bdk Budukh 20.1.3
bdm Badimaya 84.17
beb Benabena 57.8
bec Bengali (Chittagong) 9.7
bee Beembe 3.3.1
beg Begak-Ida'an 31.6.1
bej Beja 7.3.1
bel Belhare 26.2.3
bem Bemba 3.3.1

ben Bengali 9.7
beo Beothuk 90
ber Berta 6.1
bet Bété 3.8
bez Bezhta 20.1.1
bfd Biafada 3.2.1
bfg Berber (Figuig) 7.1
bfi Bafia 3.3.1
bga Benga 3.3.1
bgg Banggai 31.6.9
bgi Bagri 9.7
bgl Buglere 135.3
bgn Bugun 26.2.9
bgo Bongo 6.2.1
bgr Bagiro 6.2.1
bgs Baga Sitemu 3.2.2
bgv Bagvalal 20.1.1
bhi Bhili 9.7
bhn Bahinemo 52.3
bho Bhojpuri 9.7
bhu Bhumij 29.2
bia Bira 3.3.1
bid Bidiya 7.2.2
big Binga 6.2.1
bii Biri 84.17
bik Biak 31.3.2
bil Bilua 82
bim Bima 31.2
bin Binandere 57.4
biq Bilaan 31.6.7
bir Birom 3.3.9
bis Bisa 3.10.1
bit Biatah 31.6.1
biu Bisu 26.2.4
bjs Bajau (Semporna) 31.6.6
bka Baka (in Sudan) 6.2.1
bkd Binukid 31.6.8
bki Bakairí 175
bkl Bikol 31.6.3
bkn Bakundu 3.3.1
bkr Batak (Karo) 31.6.10
bkt Brokskat 9.7
bku Bakueri 3.3.1
bla Blackfoot 89.1
blc Baluchi 9.8
ble Baule 3.9
blg Balangao 31.6.4
blj Baale 6.3.8
blk Balantak 31.6.9
bln Bilin 7.3.2
blq Bole 7.2.4
blr Belorussian 9.10
blt Balti 26.2.3
blu Bena-Lulua 3.3.1
blx Biloxi 104
blz Balanta 3.2.1
bma Berber (Middle Atlas) 7.1
bmb Bimoba 3.5
bmn Bamun 3.3.1
bmr Burum 57.10
bmz Berber (Mzab) 7.1
bna Banawá 184
bnb Bunuba 84.1
bnd Bandi 3.10.2
bng Baining 78
bni Bini 3.3.4
bnj Bandjalang 84.17
bnk Bankon 3.3.1
bnl Banggarla 84.17
bnm Binumarien 57.8
bnn Banoni 31.3.1
bno Barasano (Northern) 145
bnq Beng 3.10.1
bnr Bilinara 84.17
bnt Bantik 31.6.9
bnu Bularnu 84.17
bnw Baniwa 174
boa Boazi 59.1
bob Bobangi 3.3.1
bod Bodo 26.2.2
boi Boiken 52.1
boj Bori 26.2.9
bok Boko 3.10.1
bol Bolia 3.3.1
bon Bongu 57.14
boq Bokar 26.2.9
bor Bora 148.2
bos Bosnian 9.10
bou Berber (Wargla) 7.1
boz Bozo (Tigemaxo) 3.10.2
bpa Bura-Pabir 7.2.1
bpb Bahnar (Plei Bong-Mang Yang) 29.1.2
bqb Basque (Bidasoa Valley) 8
bqg Basque (Gernica) 8
bqh Basque (Hondarribia) 8
bqi Basque (Basaburua and Imoz) 8
bql Basque (Lekeitio) 8
bqn Basque (Northern High Navarrese) 8
bqo Basque (Oñati) 8
bqr Basque (Roncalese) 8
bqs Basque (Sakana) 8
bqz Basque (Zeberio) 8
bra Brao 29.1.2
brb Bariba 3.5
brc Boruca 135.8
brd Bardi 84.16
bre Breton 9.4
brf Berber (Rif) 7.1
brh Brahui 25.2
bri Bribri 135.8
brj Burji 7.3.3
brk Berik 40
brl Baragaunle 26.2.3
brm Burmese 26.2.4
brn Burunge 7.3.4
bro Broken 206
brr Bororo 177.1
brs Barasano 145
bru Bru (Eastern) 29.1.3
brw Bru (Western) 29.1.3
bry Baruya 57.1
brz Birri 6.2.1
bse Berber (Ayt Seghrouchen Middle Atlas) 7.1
bsh Bushoong 3.3.1
bsi Berber (Siwa) 7.1
bsk Bashkir 11.3
bsl British Sign Language 207
bsm Bislama 206
bso Basque (Souletin) 8
bsq Basque 8
bsr Basari 3.2.1
bth Bathari 7.6
bti Betoi 151
btk Bontok 31.6.4
bto Batak (Toba) 31.6.10
bua Burarra 84.2
bub Bubi 3.3.1
bud Buduma 7.2.1
bug Bugis 31.6.9
bui Buli (in Indonesia) 31.3.2
buj Bujeba 3.3.1
buk Bukusu 3.3.1
bul Bulgarian 9.10
bum Buma 31.3.1
bun Buin 81
bur Burushaski 22
bus Busa 3.10.1
but Buriat 11.1
buu Buru 31.2
buw Bulu 3.3.1
buy Buli (in Ghana) 3.5
bvi Bali-Vitu 31.3.1
bwa Bandjalang (Waalubal) 84.17
bya Byansi 26.2.3
bys Bayso 7.3.3
byu Bandjalang (Yugumbir) 84.17
bzi Bauzi 37
cab Cabécar 135.8
cac Cacua 146
cad Caddo 103
cah Cahuilla 108.6
cai Chai 6.3.8
cak Cakchiquel 132
cam Camsá 141
can Candoshi 163
cap Capanahua 168
car Carib 175
cas Cashibo 168
cat Catio 136
cav Cavineña 186
cax Campa (Axininca) 174
cay Cayapa 140
cba Chumash (Barbareño) 123
cbo Chácobo 168
cch Chocho 124.7
cck Chickasaw 94
ccm Chinantec (Comaltepec) 124.3
cco Chasta Costa 87.1
ccp Cocopa 117.7
cct Choctaw 94
cde Carib (De'kwana) 175
cea Cree (Eastern) 89.1
ceb Cebuano 31.6.3
cec Chicomuceltec 132
cem Cèmuhî 31.3.1
cga Chaga 3.3.1
cha Chamorro 31.6.2
chb Chambri 53.4
chc Chechen 20.2
chd Chaudangsi 26.2.3
che Cherokee 91.2
chg Chang 26.2.2
chh Chaha 7.6
chi Chimariko 117.1
chj Chuj 132
chk Chukchi 14.1
chl Chehalis (Upper) 110.5
chm Chamalal 20.1.1
chn Chantyal 26.2.3
cho Chontal (Highland) 129
chp Chipewyan 87.1
chq Chinantec (Quiotepec) 124.3
chr Chrau 29.1.2
chs Chin (Siyin) 26.2.7
cht Chatino (Nopala) 124.8
chu Chulupí 198
chv Chuvash 11.3
chw Cham (Western) 31.6.10
chx Chontal (Huamelultec Oaxaca) 129
chy Chayahuita 165
chz Chaozhou 26.1
cic Chichewa 3.3.1
cil CiLuba 3.3.1
cin Chumash (Ineseño) 123
cjo Chichimeca-Jonaz 124.2
ckh Cheke Holo 31.3.1
ckl Chinook (Lower) 113.1
ckr Canela-Krahô 177.2
cku Chinook (Upper) 113.1
cla Clallam 110.2
clc Colac 84.17
cld Chaldean (Modern) 7.6
cle Chinantec (Lealao) 124.3
cln Cholon 167
cly Chulym 11.3
cma Chimila 137
cmc Comecrudo 102
cme Cham (Eastern) 31.6.10
cmh Chemehuevi 108.5
cmk Chemakum 111
cml Camling 26.2.3
cmn Comanche 108.5
cmr Chin (Mara) 26.2.7
cmx Comox 110.2
cmy Chontal Maya 132
cnl Canela 177.2
cnm Canamarí 185
cnt Cantonese 26.1
coa Coahuilteco 101
coc Cocama 176.6
coe Coeur d'Alene 110.3
cof Cofán 155
cog Cogui 135.1
coi Chortí 132
col Chol 132
com Comorian 3.3.1
coo Coos (Hanis) 114.2
cop Coptic 7.4
cor Cora 108.3
cos Costanoan 113.7.1
cpa Campa Pajonal Asheninca 174
cpl Chinantec (Palantla) 124.3
cpn Chepang 26.2.3
cpw Chippewa (Red Lake and Pillager) 89.1

cqt Chiquito 195
cre Cree (Plains) 89.1
crg Chiriguano 176.6
crh Chru 31.6.10
cri Crimean Tatar 11.3
crj Carijona 175
crk Creek 94
crl Carolinian 31.3.1
crn Cornish 9.4
cro Crow 104
crp Carapana 145
crq Carrier 87.1
crt Chorote 198
csc Chinantec (Sochiapan) 124.3
csf Chinantec (San Felipe Usila) 124.3
csh Cashinahua 168
csl Chinese Sign Language 207
cso Chatino (Sierra Occidental) 124.8
ctc Cuicatec 124.4
cte Chinantec (Tepetotutla) 124.3
cti Chin (Tiddim) 26.2.7
ctl Catalan 9.9
ctm Chitimacha 98
ctt Chatino (Tataltepec) 124.8
ctw Catawba 104
cua Cua 29.1.2
cub Cubeo 145
cuc Cuica 150
cui Cuiba 142
cul Culina 184
cum Chumburung 3.9
cup Cupeño 108.6
cur Curripaco 174
cut Cuitlatec 127
cuu Chuukese 31.3.1
cvc Chavacano 206
cve Chuave 57.5
cwe Columbia-Wenatchi 110.3
cya Chatino (Yaitepec) 124.8
cyg Cayuga 91.1
cyn Cheyenne 89.1
cyv Cayuvava 188
cze Czech 9.10
daa Da'a 31.6.9
dab Daba 7.2.1
dad Dadibi 71.2
dag Daga 74
dah Dahalo 7.3.4
dak Dakota 104
dam Damana 135.1
dan Dan 3.10.1
dar Darai 9.7
dat Datooga 6.3.5
daw Dâw 173
day Day 3.1
dbd Dabida 3.3.1
dbr Dutch (Brabantic) 9.5
dbw Dan (Blowo) 3.10.1
dca Dumagat (Casiguran) 31.6.4
dda Dhuwal (Dätiwuy) 84.17
ddf Daju (Dar Fur) 6.3.1
def Defaka 3.6
deg Degema 3.3.4
den Dení 184
der Dera 47
des Desano 145
det Deti 1.1
deu Deuri 26.2.2
dga Dagaare 3.5
dgb Dagbani 3.5
dge Deutsche Gebärdensprache 207
dgi Dogri 9.7
dgo Dongo 3.1
dgr Dagur 11.1
dgx Degexit'an 87.1
dha Dhaasanac 7.3.3
dhb Dharumbal 84.17
dhi Dhivehi 9.7
dhm Dhimal 26.2.3
dhr Dhargari 84.17
dhu Dhurga 84.17
dhw Dharawal 84.17
did Didinga 6.3.8
die Diegueño (Mesa Grande) 117.7
dig Digaro 26.2.9
dim Dime 7.5
din Dinka 6.3.5
dio Diola-Fogny 3.2.1
diy Diyari 84.17
diz Dizi 7.5
dja Djabugay 84.17
dji Djingili 84.20
djm Djambarrpuyngu 84.17
djn Djinang 84.17
djp Djapu 84.17
djr Djaru 84.17
dli Dutch (Limburg) 9.5
dma Duma 3.3.1
dmi Dumi 26.2.3
dmk Dumaki 9.7
dms Dimasa 26.2.2
dng Ding 3.3.1
dni Dani (Lower Grand Valley) 57.6
dnw Dangaléat (Western) 7.2.2
dob Dobel 31.2
dog Dogon 3.4
dok Doko 3.3.1
dol Dolgan 11.3
dom Domari 9.7
don Dong 28.2
dou Doutai 38
doy Doyayo 3.1
dre Drehu 31.3.1
drg Dargwa 20.1.2
dri Dari 9.8
drm Darmiya 26.2.3
dsh Danish 9.5
dts Dogon (Toro So) 3.4
dua Duala 3.3.1
dug Dullay (Gollango) 7.3.3
duk Duka 3.3.7
dul Dulong 26.2.11
dum Dumo 45.3
dun Duna 57.7
dut Dutch 9.5
duz Dutch (Zeeuws) 9.5
dyi Dyirbal 84.17
dym Dyimini 3.5
dyu Dyula 3.10.2
ebi Ebira 3.3.8
edo Edolo 66
efi Efik 3.3.2
ega Ega 3.9
egn Engenni 3.3.4
eip Eipo 57.17
eja Ejagham 3.3.1
eka Ekari 57.20
eko Ekoti 3.3.1
els Elseng 43
emb Embera 136
emc Embera Chami 136
eml Embaloh 31.6.10
emm Emmi 84.3.6
ena Enga 57.9
ene Enets 10.2
eng English 9.5
eno Enggano 31.6.10
eny Enya 3.3.1
epe Epena Pedee 136
err Erromangan 31.3.1
ese Ese Ejja 186
esm Esmeraldeño 156
ess Esselen 122
est Estonian 10.1.1
ets Etsako 3.3.4
eud Eudeve 108.2
eve Evenki 11.2
evn Even 11.2
ewa Ewe (Anglo) 3.9
ewe Ewe 3.9
ewo Ewondo 3.3.1
eya Eyak 87.2
far Faroese 9.5
fas Fasu 57.13
fbf Fula (Burkina Faso) 3.2.1
fea Frisian (Eastern) 9.5
fef Fe'fe' 3.3.1
fgo Fulani (Gombe) 3.2.1
fgu Fula (Guinean) 3.2.1
fij Fijian 31.3.1
fin Finnish 10.1.1
fio Fiote 3.3.1
fiw Fijian (Wayan) 31.3.1
fli Ful (Liptako) 3.2.1
fma Fula (Mauritanian) 3.2.1
fni Fula (Nigerian) 3.2.1
fno Frisian (North) 9.5
foe Foe 57.13
fon Fongbe 3.9
for Fore 57.8
fox Fox 89.1
frd Fordata 31.2
fre French 9.9
fri Frisian 9.5
frw Frisian (Western) 9.5
fsl Finnish Sign Language 207
fua Fula (Cameroonian) 3.2.1
fue Futuna (East) 31.3.1
ful Fulniô 177.8
fum Fulfulde (Maasina) 3.2.1
fur Fur 6.4
fus Fula (Senegal) 3.2.1
fut Futuna-Aniwa 31.3.1
fuz Fuzhou 26.1
fye Fyem 3.3.9
ga Gã 3.9
gaa Gaagudju 84.5
gad Gade 3.3.8
gae Gaelic (Scots) 9.4
gag Gagauz 11.3
gah Gahuku 57.8
gal Gallong 26.2.9
gam Gamo 7.5
gan Great Andamanese 30.1
gap Gapapaiwa 31.3.1
gar Garo 26.2.2
gau German (Upper Austrian) 9.5
gav Gavião 176.2
gay Gayo 31.6.10
gba German (Bavarian) 9.5
gbb Gbeya Bossangoa 3.1
gbc Guinea Bissau Crioulo 206
gbe German (Bern) 9.5
gbk Gbaya Kara 3.1
gbl German (Berlin) 9.5
gcy Greek (Cypriot) 9.6
gdb Gadaba 29.2
gdf Guduf 7.2.1
gdi Godié 3.8
gdl Guadeloupe Creole 206
gdr Gidar 7.2.1
gds Gadsup 57.8
gel Gela 31.3.1
geo Georgian 21
ger German 9.5
gfr Guianese French Creole 206
ggu Gureng Gureng 84.17
gha German (Hannover) 9.5
ghb Guahibo 142
gho Ghotuo 3.3.4
ghr Gahri 26.2.3
gid Gidabal 84.17
gil Gilaki 9.8
gim Gimira 7.5
gir Gula Iro 3.1
git Gitksan 113.6
giz Giziga 7.2.1
gji Gurindji 84.17
gjj Guajajara 176.6
gku Gününa Küne 204.2
gla Gelao 28.1
glc Galician 9.9
gll Galela 32.3
gln Golin 57.5
glp Gaalpu 84.17
gma German (Mansfeldisch) 9.5
gmb Guambiano 140
gml Gamilaraay 84.17
gmt Gumatj 84.17
gmw Gumawana 31.3.1
gmz Gumuz 6.5
gna Guana 197
gnb Gunbalang 84.7.2
gnd Ga'anda 7.2.1
gnn Gunin 84.21

gno Guanano 145
gny Gunya 84.17
goa Goajiro 174
god Godoberi 20.1.1
goe Goemai 7.2.4
gog Gogodala 67
goj Gojri 9.7
gok Gokana 3.3.2
gol Gola 3.2.2
gon Gondi 25.3
goo Gooniyandi 84.1
gor Gorowa 7.3.4
gos German (Ostschweiz) 9.5
gpz German (Appenzell) 9.5
grb Grebo 3.8
gre Greenlandic (East) 86
grf Garífuna 174
grg Gurr-goni 84.2
grj Guarijío 108.7
grk Greek (Modern) 9.6
grm Gurma 3.5
grn Gurenne 3.5
gro Guro 3.10.1
grp German (Ripuarian) 9.5
grr Garrwa 84.6
grs Garus 57.14
grt Gorontalo 31.6.9
grw Greenlandic (West) 86
gsl Greek Sign Language 207
gso Greenlandic (South) 86
gtg German (Thurgau) 9.5
gth German (Thuringian) 9.5
gti German (Timisoara) 9.5
gto Guató 177.3
gua Guaraní 176.6
gud Gude 7.2.1
gue Guere 3.8
gug Gugada 84.17
guh Guhu-Samane 57.4
guj Gujarati 9.7
gul Gula (in Central African Republic) 6.2.1
gum Gumbaynggir 84.17
gun Gunu 3.3.1
guq Guaque 175
gur Gurung 26.2.3
gus Gusii 3.3.1
gut Gurma (Togo) 3.5
guu Guugu Yimidhirr 84.17
gvi German (Viennese) 9.5
gwa Gwari 3.3.8
gwe German (Westphalian) 9.5
gwo Gworok 3.3.9
gyb Guayabero 142
gzu German (Zurich) 9.5
had Hadza 2
hai Haida 88
hak Hakka 26.1
hal Halia 31.3.1
ham Hamtai 57.1
han Hani 26.2.4
har Haruai 56
hat Hatam 32.1
hau Hausa 7.2.4
hav Havasupai 117.7
haw Hawaiian 31.3.1
hay Hayu 26.2.3
hba Hebrew (Modern Ashenazic) 7.6
hcr Haitian Creole 206
hdi Hdi 7.2.1
heb Hebrew (Modern) 7.6
heh Hehe 3.3.1
hei Heiltsuk 109.1
hem Hemba 3.3.1
her Herero 3.3.1
hhu Hanga Hundi 52.1
hia Hianacoto 175
hid Hidatsa 104
hil Hiligaynon 31.6.3
hin Hindi 9.7
hix Hixkaryana 175
hks Hong Kong Sign Language 207
hlb Halbi 9.7
hlk Halkomelem 110.2
hln Halang 29.1.2
hlp Hualapai 117.7
hma Hmar 26.2.7
hmb Huambisa 162
hmd Hmong Daw 27
hmi Huitoto (Minica) 148.2
hmo Hmong Njua 27
hmr Hamer 7.5
hmu Huitoto (Muinane) 148.2
hnd Hunde 3.3.1
hnk Hinukh 20.1.1
hnn Hanunóo 31.6.3
hno Haida (Northern) 88
ho Ho 29.2
hoa Hoava 31.3.1
hok Hokkien 26.1
hol Holoholo 3.3.1
hop Hopi 108.4
hpd Hupda 173
hre Hre 29.1.2
hrr Harari 7.6
hrs Harsusi 7.6
hsl Haisla 109.1
htc Huastec 132
htt Huitoto 148.2
hua Hua 57.8
hui Huichol 108.3
hum Huitoto (Murui) 148.2
hun Hungarian 10.1.2
hup Hupa 87.1
hve Huave (San Mateo del Mar) 131
hwc Hawaiian Creole 206
hwr Hawrami 9.8
hya Haya 3.3.1
hyo Hyow 26.2.7
hzb Hunzib 20.1.1
iaa Iaai 31.3.1
iat Iatmul 52.1
iau Iau 38
iba Iban 31.6.10
ibi Ibibio 3.3.2
ibn Ibanag 31.6.4
ice Icelandic 9.5
ics Icelandic Sign Language 207
idn Iduna 31.3.1
ido Idoma 3.3.5
idu Idu 26.2.9
ifi Italian (Fiorentino) 9.9
ifm Ifumu 3.3.1
ifu Ifugao (Batad) 31.6.4
iga Inga 169
igb Igbo 3.3.6
ige Igede 3.3.5
ign Ignaciano 174
igs Ingessana 6.3.2
iha Iha 34
iir Indonesian (Irianese) 31.6.10
ijo Ijo (Kolokuma) 3.6
ik Ik 6.3.3
ika Ika 135.1
ila Ila 3.3.1
ill Illinois 89.1
ilo Ilocano 31.6.4
imo Imonda 46
ina Inanwatan 59.2
ind Indonesian 31.6.10
ing Ingush 20.2
inj Indonesian (Jakarta) 31.6.10
inr Inuktitut (Rankin Inlet) 86
ins Inuktitut (Salluit) 86
inu Iñupiaq 86
ipi Indo-Pakistani Sign Language (Indian dialects) 207
ipk Indo-Pakistani Sign Language (Karachi dialect) 207
iql Inuktitut (Quebec-Labrador) 86
iqu Iquito 158
ird Irish (Donegal) 9.4
iri Irish 9.4
irm Irish (Munster) 9.4
irq Iraqw 7.3.4
irr Irarutu 31.3.2
irs Irish Sign Language 207
irx Iranxe 174
isa I'saka 45.2
ise Isekiri 3.3.3
ish Ishkashmi 9.8
isi Isirawa 39
isl International Sign 207
isn Isnag 31.6.4
iso Isoko 3.3.4
iss Israeli Sign Language 207
ita Italian 9.9
itb Italian (Bologna) 9.9
ite Itelmen 14.2
itg Italian (Genoa) 9.9
itn Italian (Napolitanian) 9.9
ito Itonama 187
itu Italian (Turinese) 9.9
itw Itawis 31.6.4
itz Itzaj 132
iva Ivatan 31.6.4
ivs Ivatan (Southern) 31.6.4
iwa Iwaidja 84.8
iwm Iwam 52.5
ixc Ixcatec 124.7
ixi Ixil 132
izh Izhor 10.1.1
izi Izi 3.3.6
jab Jabêm 31.3.1
jad Jad 26.2.3
jah Jahai 29.1.1
jak Jakaltek 132
jam Jaminjung 84.9
jaq Jaqaru 170
jar Jarawa (in Nigeria) 3.3.1
jav Javanese 31.6.10
jbt Jabutí 181
jcr Jamaican Creole 206
jeb Jebero 165
jeh Jeh 29.1.2
jel Jeli 3.10.2
jem Jemez 105
jia Jiarong 26.2.13
jib Jibbali 7.6
jin Jino 26.2.4
jiv Jivaro 162
jlu Jur Luwo 6.3.5
jmm Jamamadi 184
jmo Jur Mödö 6.2.1
jng Jingpho 26.2.5
joh Johari 9.7
jom Jomang 3.7
jpn Japanese 17
jpr Japreria 175
jrn Juruna 176.6
jrw Jarawa (in Andamans) 30.2
jua Juat 84.17
jug Jugli 26.2.2
juh Ju|'hoan 1.2
juk Jukun 3.3.9
jum Júma 176.5
jun Juang 29.2
jur Jurchen 11.2
jva Javaé 177.4
jwr Jarawara 184
kaa Karó (Arára) 176.4
kab Kabardian 19
kac Kachari 26.2.2
kad Kadugli 5
kae Kaki Ae 72
kag Kayu Agung 31.6.10
kai Kaian 53.3
kaj Kaure 44
kak Kamano-Kafe 57.8
kal Kalami 9.7
kam Kambera 31.2
kan Kana 3.3.2
kao Karao 31.6.4
kap Kela (Apoze) 31.3.1
kaq Kaurna 84.17
kar Kara (in Central African Republic) 6.2.1
kas Kashmiri 9.7
kat Kâte 57.10
kau Kaulong 31.3.1
kaw Kaiwá 176.6
kay Kayardild 84.18
kaz Kazakh 11.3
kba Kamba 3.3.1
kbi Kabui 26.2.7
kbl Kabyle 7.1
kbo Kambot 53.2

kbr Kayan (Baram) 31.6.1
kbt Kabatei 9.8
kbu Kanembu 6.9
kbw Karen (Bwe) 26.2.6
kby Kabiyé 3.5
kch Koyra Chiini 6.11
kda Konda 25.3
kdg Karipuna do Guapore 168
kdw Kadiwéu 199
kdz Kadazan 31.6.1
kea Kanjobal (Eastern) 132
ked Kedang 31.2
kef Kefa 7.5
kei Kei 31.2
kek K'ekchí 132
kel Kele 31.3.1
kem Kemant 7.3.2
ken Kenga 6.2.1
keo Ke'o 31.2
ker Kera 7.2.2
ket Ket 13
keu Kenyah (Uma' Lung) 31.6.1
kew Kewa 57.9
kfc Kriol (Fitzroy Crossing) 206
kfe Koromfe 3.5
kfy Kirghiz (Fu-Yu) 11.3
kga Kinga 3.3.1
kgi Konyagi 3.2.1
kgl Kaugel 57.5
kgm Kagoma 3.3.9
kgr Kaguru 3.3.1
kgt Kangiryuarmiut 86
kgu Kalkatungu 84.17
kgz Kirghiz 11.3
kha Khalkha 11.1
khg Khaling 26.2.3
khi Khinalug 20.1.3
khk Khakas 11.3
khl Khalaj 11.3
khm Khmer 29.1.5
kho Khoekhoe 1.1
khr Kharia 29.2
khs Khasi 29.1.4
khu Khumi 26.2.7
khv Khvarshi 20.1.1
khw Khowar 9.7
kic Kickapoo 89.1
kie Kire 53.5
kij Kitja 84.4
kik Kikuyu 3.3.1
kil Kiluba 3.3.1
kim Kimaghama 60
kin Kinyarwanda 3.3.1
kio Kiowa 105
kiq Kalmyk (Issyk-Kul) 11.1
kir Kirma 3.5
kis Kisi 3.2.2
kit Kitsai 103
kiu Kui (in Indonesia) 57.19
kiw Kiwai 68
kje Koreguaje 145
kji Kurmanji 9.8
kjn Kunjen 84.17
kjo Konjo 31.6.9
kjr Koyra 7.5
kkb Konkomba 3.5
kkl Kata Kolok 207
kkn Konkani 9.7
kko Koranko 3.10.2
kkp Karakalpak 11.3
kkr Kirikiri 38
kkt Kokota 31.3.1
kku Korku 29.2
kkv Kaliai-Kove 31.3.1
kkw Karankawa 100
kky Kankanay 31.6.4
kkz Kokni 9.7
kla Klao 3.8
klb Kilba 7.2.1
klg Kulung 26.2.3
klh Kalasha 9.7
kli Kaili 31.6.9
klk Koh (Lakka) 3.1
kll Kaluli 66
klm Klamath 113.2
kln Kolana 57.19
klp Kalapuya 116
klq Kalam 57.14
klr Koluri 9.8
kls Kalispel 110.3
klt Kelabit 31.6.1
klv Kilivila 31.3.1
klw Kiliwa 117.7
kly Kala Lagaw Ya 84.17
klz Kalanga 3.3.1
kma Kamaiurá 176.6
kmb Kombai 57.3
kmh Kham 26.2.3
kmi Kami 3.3.1
kmj Karimojong 6.3.5
kmk Kalmyk 11.1
kmn Kuman 57.5
kmo Koiali (Mountain) 57.12
kmp Kunimaipa 57.11
kmr Kamoro 57.2
kms Kamass 10.2
kmt Kemtuik 41
kmu Khmu' 29.1.8
kmw Kamu 84.3.2
kna Karitiâna 176.1
knb Kanum (Bädi) 63
knc Kugu Nganhcara 84.17
knd Kannada 25.4
kng Kaingang 177.2
kni Konni 3.5
knk Kanakuru 7.2.4
knm Kunama 6.7
knn Kinnauri 26.2.3
knp Kanum (Ngkâlmpw) 63
knq Kriol (Ngukurr) 206
knr Kanuri 6.9
knu Konua 80
knw Konkow 113.3
koa Koasati 94
kob Kobon 57.14
kod Kodava 25.4
koe Koegu 6.3.8
koh Kohumono 3.3.2
koi Koiari 57.12
kok Kokborok 26.2.2
kol Kolami 25.1
kom Komo 6.6
kon Kongo 3.3.1
koo Kola 31.2
kop Komi-Permyak 10.1.1
kor Korean 18
kos Kosraean 31.3.1
kot Kota 25.4
kou Kom 3.3.1
kow Ko (Winye) 3.5
koy Koya 25.3
kpa Kpan 3.3.9
kpe Kpelle 3.10.2
kpm Kapampangan 31.6.4
kpn Kapingamarangi 31.3.1
kpo Kposo 3.9
kpw Karen (Pwo) 26.2.6
kra Kara (in Papua New Guinea) 31.3.1
krb Kiribati 31.3.1
krc Karachay-Balkar 11.3
krd Kurdish (Central) 9.8
kre Kresh 6.2.2
krf Korafe 57.4
krg Karanga 3.3.1
krh Krahô 177.2
kri Kariri 177.5
krj Karadjeri 84.17
krk Karok 118
krl Karelian 10.1.1
krm Karaim 11.3
krn Korana 1.1
kro Krongo 5
krq Kerek 14.1
krr Kairiru 31.3.1
krt Karata 20.1.1
krw Korowai 57.3
kry Koryak 14.1
krz Kryz 20.1.3
ksa Keresan (Santa Ana) 106
kse Koyraboro Senni 6.11
ksg Karen (Sgaw) 26.2.6
ksh Kashaya 117.3
ksi Ksingmul 29.1.8
ksl Kenyan Sign Language 207
ksm Kasem 3.5
ksp Kosop 3.1
ksr Kisar 31.2
kss Kisi (Southern) 3.2.2
ksu Kashubian 9.10
kta Koita 57.12
ktb Kituba 206
ktc Katcha 5
kte Kete 3.3.1
kth Kutchin 87.1
kti Kati (in West Papua, Indonesia) 57.18
ktk Kotoko 7.2.1
ktl Katla 3.7
ktn Ketengban 57.17
kto Kato 87.1
ktp Ketapang 31.6.10
ktt Kott 13
ktu Katu 29.1.3
kty Khanty 10.1.2
ktz Kati (in Afghanistan) 9.7
kua Kualan 31.6.10
kug Kunming 26.1
kui Kui (in India) 25.3
kuk Kukú 6.3.5
kul Kullo 7.5
kum Kumauni 9.7
kun Kuna 135.4
kuo Kuot 79
kuq Kumyk 11.3
kur Kurukh 25.2
kus Kusunda 24
kut Kutenai 112
kuu Kuuku Ya'u 84.17
kuv Kuvi 25.3
kuy Kutai 31.6.10
kuz Kulamanen 31.6.8
kwa Kwaio 31.3.1
kwb Kwerba 39
kwe Kanjobal (Western) 132
kwk Kwakw'ala 109.1
kwm Kwami 7.2.4
kwn Kwangali 3.3.1
kwo Kwoma 52.1
kwr Kwamera 31.3.1
kws Kawaiisu 108.5
kwz Kwazá 180
kxo Kxoe 1.1
kya Kuku-Yalanji 84.17
kyg Kenyang 3.3.1
kyk Kyaka 57.9
kyl Kayah Li (Eastern) 26.2.6
kym Krymchak 11.3
kyn Koyukon 87.1
kyo Kanyok 3.3.1
kyp Kayapó 177.2
kyq Kyuquot 109.2
kyr Karkar-Yuri 48
kzh Kam (Zhanglu) 28.2
kzy Komi-Zyrian 10.1.1
laa Laal 4
lac Lacandón 132
lad Ladakhi 26.2.3
laf Lafofa 3.7
lag Lagwan 7.2.1
lah Lahu 26.2.4
lai Lai 26.2.7
lak Lak 20.1.2
lal Lalo 26.2.4
lam Lamé 7.2.3
lan Lango 6.3.5
lao Lao 28.2
lar Laragia 84.10
lat Latvian 9.3
lau Lau 31.3.1
lav Lavukaleve 82
laz Laz 21
lbu Lunda 3.3.1
lch Lachi 28.1
lcr Lesser Antillean French Creole 206
lda Luganda 3.3.1
ldn Ladin 9.9

ldo Londo 3.3.1
ldu Lendu 6.2.3
leb Lebeo 3.3.1
leg Lega 3.3.1
lel Lele 7.2.2
len Lenakel 31.3.1
lep Lepcha 26.2.8
les Lese 6.2.5
let Leti 31.2
lew Lewo 31.3.1
lez Lezgian 20.1.3
lga Luangiua 31.3.1
lge Low German 9.5
lgh Lughat al-Isharat al-Lubnaniya 207
lgi Langi 3.3.1
lgp Língua Gestual Portuguesa 207
lgt Logoti 6.2.6
lgu Longgu 31.3.1
lha Laha 28.1
lho Lhomi 26.2.3
lii Lingua Italiana dei Segni 207
lil Lillooet 110.3
lim Limbu 26.2.3
lin Lingala 3.3.1
lis Lisu 26.2.4
lit Lithuanian 9.3
liv Liv 10.1.1
lje Lauje 31.6.9
lkk Lakkia 28.1
lkt Lakhota 104
lla Lamu-Lamu 84.17
llm Lelemi 3.9
lma Loma 3.10.2
lmb Lamba 3.3.1
lmg Lamang 7.2.1
lmh Lamaholot 31.2
lml Limilngan 84.11
lmn Lamani 9.7
lmp Lampung 31.6.10
lnd Linda 3.1
lng Lengua 197
lno Ladino 9.9
lns Lamnso' 3.3.1
lnw Lonwolwol 31.3.1
lob Lobi 3.5
lok Loko 3.10.2
lon Loniu 31.3.1
lot Lotha 26.2.7
lou Lou 31.3.1
lov Loven 29.1.2
loz Lozi 3.3.1
lrd Lardil 84.17
lrk Larike 31.2
lsa Lengua de Señas Argentina 207
lsb Língua de Sinais Brasileira 207
lse Lengua de Señas Española 207
lsf Langue des Signes Française 207
lsq Langue des Signes Québecoise 207
lu Lü 28.2
lua Lua 3.1
luc Lucazi 3.3.1
lud Lun Dayeh 31.6.1
lug Lugbara 6.2.6
lui Luiseño 108.6
lul Lule 200
lum Lummi 110.2
lun Lungchang 26.2.2
luo Luo 6.3.5
lur Luri 9.8
lus Lushootseed 110.2
luv Luvale 3.3.1
lux Luxemburgeois 9.5
luy Luyia 3.3.1
lye Lyele 3.5
ma Ma 3.1
maa Maasai 6.3.5
mab Maba 6.8
mac Macushi 175
mad Ma'di 6.2.6
mae Mae 31.3.1
mag Magar 26.2.3
mah Mari (Hill) 10.1.1
mai Maithili 9.7
maj Majang 6.3.8
mak Makah 109.2
mal Malagasy 31.6.1
mam Mam 132
man Mano 3.10.1
mao Maori 31.3.1
map Mapudungun 202
mar Maricopa 117.7
mas Masa 7.2.3
mau Maung 84.8
mav Maring 57.5
maw Maninka (Western) 3.10.2
max Maxakalí 177.6
may Maybrat 32.4
maz Mazahua 124.5
mba Mba 3.1
mbb Mbabaram 84.17
mbe Mbere 3.3.1
mbg Mbugu 7.3.4
mbi Mbili 3.3.1
mbl Mbole 3.3.1
mbm Mbum 3.1
mbo Monumbo 55
mbr Mbara 7.2.1
mbt Mangbetu 6.2.4
mbu Manambu 52.1
mby Mbay 6.2.1
mbz Mbe' 3.3.1
mca Maca 198
mcc Mochica 166
mcd Macedonian 9.10
mce Mískito Coast English Creole 206
mcf Michif 89.1
mcg Macaguán 142
mch Machiguenga 174
mcn Macuna 145
mco Mixe (Coatlán) 130
mcr Mauritian Creole 206
mcs Miwok (Central Sierra) 113.7.2
mcv Mocoví 199
mda Mada (in Cameroon) 7.2.1
mdb Mudburra 84.17
mde Mende 3.10.2
mdg Mundang 3.1
mdk Mandinka 3.10.2
mdl Madngele 84.3.2
mdm Madimadi 84.17
mdn Mandan 104
mdo Mbodomo 3.1
mdr Madurese 31.6.10
mds Manadonese 31.6.10
mdu Mündü 3.1
mdz Mada (in Nigeria) 3.3.9
mea Meyah 33
meb Melayu Betawi 206
mee Me'en 6.3.8
meh Mehri 7.6
mei Meithei 26.2.7
mek Mekens 176.5
mel Melanau 31.6.1
mem Manem 46
men Menomini 89.1
mer Meryam Mir 69
mey Menya 57.1
mga Mondunga 3.1
mgd Mongondow 31.6.9
mge Mnong (Eastern) 29.1.2
mgg Mangghuer 11.1
mgh Magahi 9.7
mgi Magi 57.15
mgk Mungaka 3.3.1
mgl Mingrelian 21
mgn Magindanao 31.6.8
mgo Mongo 3.3.1
mgu Musgu 7.2.1
mhc Mahican 89.1
mhi Marathi 9.7
mhk Mehek 52.4
mhm Mah Meri 29.1.1
mhu Mbalanhu 3.3.1
mic Micmac 89.1
mid Midob 6.3.6
mie Mien 27
mig Migama 7.2.2
mii Miisiirii 6.3.9
mij Miju 26.2.9
mik Mikir 26.2.7
mil Milang 26.2.9
min Minangkabau 31.6.10
mir Miriwung 84.4
mis Miskito 134
mit Mituku 3.3.1
miy Miya 7.2.4
miz Mizo 26.2.7
mja Mixtec (Jamiltepec) 124.4
mjc Mixtec (San Juan Colorado) 124.4
mjk Manjaku 3.2.1
mka Mauka 3.10.2
mkb Miwok (Bodega) 113.7.2
mkd Makonde 3.3.1
mke Mekeo 31.3.1
mkg Mandinka (Gambian) 3.10.2
mkh Mongol (Khamnigan) 11.1
mki Mikasuki 94
mkl Maklew 62
mkn Mankanya 3.2.1
mko Mokilko 7.2.2
mkr Mikarew 53.3
mks Makassar 31.6.9
mku Maranungku 84.3.6
mkw Máku 172
mkz Makaa 3.3.1
mla Mambila 3.3.1
mlc Malacca Creole 206
mle Maale 7.5
mlg Malgwa 7.2.1
mli Mali 78
mlk Malakmalak 84.3.4
mll Molala 113.4
mlm Mlabri (Minor) 29.1.8
mlt Maltese 7.6
mlu Maleu 31.3.1
mly Malay 31.6.10
mma Mandaic (Modern) 7.6
mmb Mangap-Mbula 31.3.1
mme Mari (Meadow) 10.1.1
mmn Mamanwa 31.6.3
mmo Mordvin (Moksha) 10.1.1
mmp Mampruli 3.5
mms Mamasa 31.6.9
mmu Malay (Ulu Muar) 31.6.10
mmy Mamyu 6.2.5
mmw Mambwe 3.3.1
mmz Mam (Southern) 132
mna Muna 31.6.9
mnc Manchu 11.2
mnd Mandarin 26.1
mne Maidu (Northeast) 113.3
mng Manggarai 31.2
mni Moni 57.20
mnj Mantjiltjara 84.17
mnk Maninka 3.10.2
mnm Manam 31.3.1
mno Mono (in United States) 108.5
mnr Mandar 31.6.9
mns Mansi 10.1.2
mnt Mentawai 31.6.10
mnu Mun 27
mnv Minaveha 31.3.1
mnx Manx 9.4
mny Margany 84.17
moa Mono-Alu 31.3.1
moc Moca 7.5
moe Mordvin (Erzya) 10.1.1
mof Mofu-Gudur 7.2.1
mog Moghol 11.1
moh Mohawk 91.1
moj Mojave 117.7
mok Mokilese 31.3.1
mol Moldavian 9.9
mom Mombum 61
mon Mon 29.1.6
moo Mooré 3.5
mop Mopan 132
mor Mor 31.3.2
mos Mosetén 190
mot Motuna 81
mou Moru 6.2.6
mov Movima 189
mpa Murrinh-Patha 84.3.3
mpo Mpongwe 3.3.1
mpr Maipure 174

mpu Mpur 32.2
mpy Mapoyo 175
mqc Martinique Creole 206
mra Mara 84.13
mrc Marchha 26.2.3
mrd Marind 59.1
mrg Margi 7.2.1
mrh Marrithiyel 84.3.6
mri Moraori 63
mrk Murik 53.4
mrl Murle 6.3.8
mrn Maranao 31.6.8
mro Moro 3.7
mrq Marquesan 31.3.1
mrr Maringarr 84.3.6
mrs Mairasi 57.16
mrt Martuthunira 84.17
mru Maru 26.2.4
mrw Muruwari 84.17
msc Muisca 135.2
mse Munsee 89.1
msg Mising 26.2.9
msh Marshallese 31.3.1
msk Masakin 3.7
msl Masalit 6.8
msm Musom 31.3.1
msn Maisin 31.3.1
mss Miwok (Southern Sierra) 113.7.2
msy Magar (Syangja) 26.2.3
mta Mentuh Tapuh 31.6.1
mtb Matuumbi 3.3.1
mtg Montagnais 89.1
mth Mixe (Tlahuitoltepec) 130
mti Motilón (Chibchan) 135.5
mtk Matukar 31.3.1
mtl Mattole 87.1
mto Malto 25.2
mtp Mixe (Totontepec) 130
mts Matis 168
mtt Massachusett 89.1
mtu Motu 31.3.1
mua Makua 3.3.1
mud Mundani 3.3.1
mug Mugil 57.14
muh Muher 7.6
mui Muinane 148.1
mul Mulao 28.2
mum Mumuye 3.1
mun Mundari 29.2
muo Muong 29.1.9
mup Mupun 7.2.4
mur Mursi 6.3.8
mus Mussau 31.3.1
mut Mutsun 113.7.1
muu Mundurukú 176.3
muy Muyuw 31.3.1
mwa Martu Wangka 84.17
mwb Manobo (Western Bukidnon) 31.6.8
mwc Mawchi 9.7
mwe Mwera 3.3.1
mwl Miwok (Lake) 113.7.2
mwn Miwok (Northern Sierra) 113.7.2
mwo Mwotlap 31.3.1
mwp Miwok (Plains) 113.7.2
mxa Mixtec (Atatlahuca) 124.4
mxc Mixtec (Chalcatongo) 124.4
mxg Mixtec (San Miguel el Grande) 124.4
mxj Mixtec (Jicaltepec) 124.4
mxl Mixtec (Alacatlatzala) 124.4
mxm Mixtec (Molinos) 124.4
mxo Mixtec (Ocotepec) 124.4
mxp Mixtec (Peñoles) 124.4
mxs Mixtec (Silacayoapan) 124.4
mxt Mixtec (Ayutla) 124.4
mxu Mixtec (Chayuco) 124.4
mxy Mixtec (Yosondúa) 124.4
mxz Mixtec (Coatzospan) 124.4
mya Ma'ya 31.3.2
myi Mangarrayi 84.12
myk Malay (Kuala Lumpur) 31.6.10
mym Malayalam 25.4
myn Ma'anyan 31.6.1
myo Mayo 108.2
myr Matsés 168
myy Mayi-Yapi 84.17
mzc Mazatec (Chiquihuitlán) 124.7
mzh Mazatec (Huautla) 124.7
mzn Mazanderani 9.8
naa Neo-Aramaic (Amadiya) 7.6
nab Nabak 57.10
nad Nadëb 173
naf Nafaanra 3.5
nag Nagatman 49
nah Nahali 23
nai Nanai 11.2
naj Neo-Aramaic (Arbel Jewish) 7.6
nak Nakanai 31.3.1
nal Nalik 31.3.1
nam Namia 52.6
nan Nandi 6.3.5
nap Neo-Aramaic (Persian Azerbaijan) 7.6
nar Nara (in Ethiopia) 6.3.4
nas Nasioi 81
nat Natchez 96
nau Nauruan 31.3.1
nav Navajo 87.1
nax Naxi 26.2.10
nbb Nambas (Big) 31.3.1
nbd Nubian (Dongolese) 6.3.6
nbe Ngombe 3.3.1
nbh Nubian (Hill) 6.3.6
nbk Nambakaengö 83
nbo Nyambo 3.3.1
nbr Ngäbere 135.3
nca Nicobarese (Car) 29.1.7
ndb Ndebele (in South Africa) 3.3.1
nde Nande 3.3.1
ndg Ndogo 3.1
ndi Ngbandi 3.1
ndj Ndjébbana 84.15
ndo Ndonga 3.3.1
ndt Ndut 3.2.1
ndu Ndumu 3.3.1
ndy Ndyuka 206
ned Nederlandse Gebarentaal 207
neg Negidal 11.2
neh Nehan 31.3.1
nel Nelemwa 31.3.1
nen Nenets 10.2
nep Nepali 9.7
nev Nevome 108.8
new Newari (Kathmandu) 26.2.3
nez Nez Perce 113.5
nga Nganasan 10.2
ngb Ngbaka 3.1
ngd Ngad'a 31.2
nge Ngemba 3.3.1
ngg Ngan'gityemerri 84.3.5
ngi Ngiyambaa 84.17
ngj Ngadjumaja 84.17
ngk Ngankikurungkurr 84.3.5
ngl Ngalakan 84.7.3
ngm Ngambay 6.2.1
ngn Ngandi 84.7.4
ngo Ngoni 3.3.1
ngr Ngarinyeri 84.17
ngt Naga (Tangkhul) 26.2.7
ngu Nguna 31.3.1
ngw Ngawun 84.17
ngy Ngarinyman 84.17
ngz Ngizim 7.2.4
nha Nhanda 84.17
nhc Nahuatl (Central) 108.1
nhh Nahuatl (Huasteca) 108.1
nhm Nahuatl (Michoacán) 108.1
nhn Nahuatl (North Puebla) 108.1
nhp Nahuatl (Pochutla) 108.1
nht Nahuatl (Tetelcingo) 108.1
nhu Nahuatl (Huauchinango) 108.1
nhx Nahuatl (Xalitla) 108.1
nia Nias 31.6.10
nic Nicobarese 29.1.7
nif Niuafo'ou 31.3.1
nih Nihon Shuwa (Japanese Sign Language) 207
nim Nimboran 41
nin Ningil 51.3
nis Nishi 26.2.9
nit Nitinaht 109.2
niu Niuean 31.3.1
niv Nivkh 15
nju Nyungar 84.17
nkb Ngalkbun 84.7.2
nke Nkem 3.3.1
nkk Nakkara 84.14
nkn Nkonya 3.9
nko Nkore-Kiga 3.3.1
nkr Nukuoro 31.3.1
nkt Nyah Kur (Tha Pong) 29.1.6
nku Nubian (Kunuz) 6.3.6
nlu Ngarluma 84.17
nma Naga (Mao) 26.2.7
nmb Nambikuára 179
nmd Nomad 65
nmi Nahuatl (Mecayapan Isthmus) 108.1
nmp Nahuatl (Milpa Alta) 108.1
nnc Nancowry 29.1.7
nnd Nend 57.14
nne Nengone 31.3.1
nng Nyanga 3.3.1
nnk Nankina 57.10
nnm Nanumea 31.3.1
nnn Nuni (Northern) 3.5
nnt Nanticoke 89.1
nob Nobiin 6.3.6
noc Nocte 26.2.2
nog Noghay 11.3
nok Noghay (Karagash) 11.3
nom Nomatsiguenga 174
non Noni 3.3.1
noo Noon 3.2.1
nor Norwegian 9.5
npa Nahuatl (Pajapan) 108.1
nph Nar-Phu 26.2.3
npi Nigerian Pidgin 206
npn Naga Pidgin 206
npu Napu 31.6.9
nrm Narom 31.6.1
nro Nharo 1.1
nse Nsenga 3.3.1
nsg Nisgha 113.6
nsk Naskapi 89.1
nsn Nisenan 113.3
nsy Neo-Aramaic (Assyrian) 7.6
nsz Nahuatl (Sierra de Zacapoaxtla) 108.1
nte Norsk Tegnspråk 207
nti Ngiti 6.2.3
ntj Ngaanyatjarra 84.17
ntn Nateni 3.5
nto Ntomba 3.3.1
ntu Nenets (Tundra) 10.2
nua Nuaulu 31.2
nub Nubi 206
nue Nuer 6.3.5
nug Nunggubuyu 84.7.5
nuk Nukak 146
nun Nung (in Vietnam) 28.2
nup Nupe 3.3.8
nus Nusu 26.2.4
nuu Nuuchahnulth 109.2
nvs Nivkh (South Sakhalin) 15
nwd Newari (Dolakha) 26.2.3
nya Nyawaygi 84.17
nyg Nyangi 6.3.3
nyh Nyiha 3.3.1
nyi Nyimang 6.3.7
nyk Nyamkad 26.2.3
nyl Nyelâyu 31.3.1
nym Nyamwezi 3.3.1
nyn Nyigina 84.16
nyr Nyangumarda 84.17
nyu Nyulnyul 84.16
nza Nzakara 3.1
nze Naga (Zeme) 26.2.7
nzs New Zealand Sign Language 207
obg Ogbronuagum 3.3.2
obk Obokuitai 38
obo Obolo 3.3.2
oca Ocaina 148.2
occ Occitan 9.9
ocu Ocuilteco 124.5
ogb Ogbia 3.3.2
oi Oi 29.1.2
oir Oirat 11.1
oix Otomí (Ixtenco) 124.5

oji Ojibwa (Eastern) 89.1
ojm Ojibwe (Minnesota) 89.1
ojs Ojibwa (Severn) 89.1
oka Okanagan 110.3
oko Oriya (Kotia) 9.7
oks Oksapmin 50
oku Oku 3.3.1
olm Oloh Mangtangai 31.6.1
olo Olo 51.3
olu Olutec 130
oma Omagua 176.6
omh Omaha 104
omi Ömie 57.12
ond Oneida 91.1
one One 51.4
ong Onge 30.2
onn Onondaga 91.1
ono Ono 57.10
ood O'odham 108.8
ora Ora (in Nigeria) 3.3.4
orb Oromo (Boraana) 7.3.3
orc Oroch 11.2
ord Ordos 11.1
ore Orejón 145
orh Oromo (Harar) 7.3.3
ori Orig 3.7
ork Orok 11.2
orl Orokolo 72
orm Ormuri 9.8
oro Orokaiva 57.4
orw Oromo (Waata) 7.3.3
ory Orya 40
osa Osage 104
osm Otomí (Santiago Mexquititlan) 124.5
oss Ossetic 9.8
otm Otomí (Mezquital) 124.5
oto Oto 104
otr Otoro 3.7
ots Otomí (Sierra) 124.5
owc Oromo (West-Central) 7.3.3
oya Oriya 9.7
paa Pa'a 7.2.4
pac Pacoh 29.1.3
pad Padoe 31.6.9
pae Páez 139
pag Pagu 32.3
pai Paiwan 31.4
pak Pakanha 84.17
pal Palauan 31.6.5
pam Pame 124.6
pan Panjabi 9.7
pap Papiamentu 206
par Päri 6.3.5
pat Patep 31.3.1
pau Paumarí 184
paw Pawaian 71.1
pba Pima Bajo 108.8
pcm Pocomam 132
pdp Podopa 71.2
pec Pech 135.6
pem Pemon 175
pen Pengo 25.3
per Pero 7.2.4
pga Pilagá 199
pgl Pingilapese 31.3.1
phl Phlong 26.2.6
pia Piaroa 143
pil Pileni 31.3.1
pin Pintupi 84.17
pip Pipil 108.1
pir Piro 174
pis Pisa 57.3
pit Pitjantjatjara 84.17
pkm Pokomchí 132
pkn Paakantyi 84.17
pkt Pokot 6.3.5
pla Playero 142
plb Polabian 9.10
plg Palaung 29.1.8
plh Paulohi 31.2
plk Palikur 174
plp Pulopetak 31.6.1
plr Palor 3.2.1
pmc Pomo (Central) 117.3
pme Pomo (Eastern) 117.3
pmn Pomo (Northern) 117.3
pms Paamese 31.3.1
pna Pamona 31.6.9
png Pangwa 3.3.1
pnn Pangasinan 31.6.4
pno Paiute (Northern) 108.5
pnr Panare 175
pnu Punu 27
pny Panyjima 84.17
poa Po-Ai 31.3.1
poc Pocomam Oriental 132
pod Podoko 7.2.1
pog Pogoro 3.3.1
poh Pohnpeian 31.3.1
pok Poko-Rawo 45.1
pol Polish 9.10
pop Popoloca (Metzontla) 124.7
por Portuguese 9.9
pot Potawatomi 89.1
pow Powhatan 89.1
ppc Piapoco 174
ppi Pitta Pitta 84.17
pra Prasuni 9.7
prc Parecis 174
prd Parji (Dravidian) 25.1
pre Pare 3.3.1
prh Pirahã 183
pri Príncipense 206
prk Purki 26.2.3
pro Provençal 9.9
prs Persian 9.8
prt Piratapuyo 145
psh Pashto 9.8
psj Popoloca (San Juan Atzingo) 124.7
psl Plains-Indians Sign Language 207
psm Passamaquoddy-Maliseet 89.1
pso Pomo (Southeastern) 117.3
psv Popoloca (San Vicente Coyotepec) 124.7
pta Paita 31.3.1
pte Paite 26.2.7
ptp Patpatar 31.3.1
ptt Pattani 26.2.3
ptw Patwin 113.8
pui Puinave 144
puk Parauk 29.1.8
pul Puluwat 31.3.1
pum Pumi 26.2.12
pun Pungupungu 84.3.1
puq Puquina 192
pur Purépecha 125
put Paiute (Southern) 108.5
pwn Pawnee 103
qaf Qafar 7.3.3
qan Quechua (Ancash) 169
qaw Qawasqar 203
qay Quechua (Ayacucho) 169
qbo Quechua (Bolivian) 169
qca Quechua (Cajamarca) 169
qch Quiché 132
qco Quechua (Cochabamba) 169
qcu Quechua (Cuzco) 169
qec Quechua (Ecuadorean) 169
qhu Quechua (Huallaga) 169
qia Qiang 26.2.12
qim Quechua (Imbabura) 169
qta Quechua (Tarma) 169
que Quechan 117.7
qui Quileute 111
rad Rade 31.6.10
rag Raga 31.3.1
ram Rama 135.7
rao Rao 53.1
rap Rapanui 31.3.1
ras Rashad 3.7
rav Romani (Ajia Varvara) 9.7
raw Rawang 26.2.11
rbg Romani (Burgenland) 9.7
rbu Romani (Bugurdzi) 9.7
rcp Russian-Chinese Pidgin (Birobidjan) 206
rej Rejang 31.6.10
rem Remo 29.2
ren Rendille 7.3.3
res Resígaro 174
ret Retuarã 145
rga Ronga 3.3.1
rgn Roglai (Northern) 31.6.10
ria Riantana 60
rik Rikbaktsa 177.7
rim Rimi 3.3.1
rit Ritharngu 84.17
rka Romani (Kalderash) 9.7
rlo Romani (Lovari) 9.7
rmb Rembarnga 84.7.6
rmc Romansch 9.9
rnd Rundi 3.3.1
rng Rengao 29.1.2
rnn Rennellese 31.3.1
rnr Romani (North Russian) 9.7
rny Runyankore 3.3.1
rom Romanian 9.9
ron Ron 7.2.4
ror Roro 31.3.1
rot Rotuman 31.3.1
rov Roviana 31.3.1
rru Runyoro-Rutooro 3.3.1
rsc Romansch (Scharans) 9.9
rse Romani (Sepecides) 9.7
rsh Rushan 9.8
rsl Russian Sign Language 207
rsm Romansch (Surmeiran) 9.9
rsu Romansch (Sursilvan) 9.9
rti Roti 31.2
rtk Rotokas 80
ruk Rukai 31.5
rum Rumu 70
run Runga 6.8
rus Russian 9.10
rut Rutul 20.1.3
rwa Rawa 57.10
rwe Romani (Welsh) 9.7
saa Sa'a 31.3.1
sab Sa'ban 31.6.1
sae Saek 28.2
sah Sahu 32.3
sak Sakao 31.3.1
sal Salinan 121
sam Samoan 31.3.1
san Sango 3.1
sao Saho 7.3.3
sap Sapuan 29.1.2
sar Sare 52.3
sav Savi 9.7
saw Sawu 31.2
say Sayultec 130
sba Sáliba (in Colombia) 143
sbg Sama (Balangingi) 31.6.6
sch Saanich 110.2
scr Serbian-Croatian 9.10
scs Saami (Central-South) 10.1.1
sdh Sindhi 9.7
sdw Sandawe 1.3
sea Southeast Ambrym 31.3.1
seb Sebei 6.3.5
sec Secoya 145
sed Sedang 29.1.2
see Seediq 31.1
sel Selknam 204.1
sem Sema 26.2.7
sen Sena 3.3.1
ser Seri 117.4
ses Sesotho 3.3.1
sey Seychelles Creole 206
sgb Sougb 33
sge Songe 3.3.1
sgl Sengele 3.3.1
sgr Sangir 31.6.9
sgs Songish 110.2
sgu Sangu 3.3.1
sha Shan 28.2
shb Shabo 6.10
shd Sherdukpen 26.2.9
she Sherpa 26.2.3
shh Sharanahua 168
shi Shiriana 154
shk Shipibo-Konibo 168
shl Shilluk 6.3.5
shm Shambala 3.3.1
shn Shona 3.3.1
sho Shoshone 108.5
shp Sahaptin (Northern) 113.5

shr Shor 11.3
shs Shasta 117.5
sht Shatt 6.3.1
shu Shuswap 110.3
shw Shawnee 89.1
shy Shira Yughur 11.1
sia Siane 57.8
sik Sika 31.2
sil Sila 6.3.1
sim Simeulue 31.6.10
sin Siona 145
sio Sio 31.3.1
sir Siar 31.3.1
sis Sisiqa 31.3.1
siu Siuslaw 114.3
ski Saami (Kildin) 10.1.1
skk Sikkimese 26.2.3
skl Sikule 31.6.10
skm Sukuma 3.3.1
skn Sikuani 142
sko Sko 45.3
skp Selkup 10.2
skr Sikaritai 38
skw Shekhawati 9.7
sla Slave 87.1
slb Saliba (in Papua New Guinea) 31.3.1
sle Samba Leko 3.1
slg Sulung 26.2.9
slo Slovene 9.10
slp Selepet 57.10
slr Salar 11.3
slv Slavey 87.1
sly Selayar 31.6.9
smd Semandang 31.6.1
sme Seme 3.8
smj Semai 29.1.1
sml Semelai 29.1.1
smn Seminole 94
smt Sahaptin (Umatilla) 113.5
sna Shina 9.7
snc Seneca 91.1
snd Senadi 3.5
sng Sinaugoro 31.3.1
snh Sinhala 9.7
snm Sanuma 154
snn Soninke 3.10.2
sno Saami (Northern) 10.1.1
snt Sentani 42
so So 6.3.3
sob Sobei 31.3.1
sod Soddo 7.6
sol Solon 11.2
som Somali 7.3.3
son Sonsorol-Tobi 31.3.1
soo Sooke 110.2
soq Soqotri 7.6
sor Sora 29.2
sou Sorbian (Upper) 9.10
spa Spanish 9.9
spc Spanish (Canary Islands) 9.9
spi Spitian 26.2.3
spo Spokane 110.3
sps Salish (Southern Puget Sound) 110.2
squ Squamish 110.2
sra Sranan 206
srb Sorbian 9.10
src Sarcee 87.1
srd Sardinian 9.9
sre Sre 29.1.2
sri Siriano 145
srl Sorbian (Lower) 9.10
srm Saramaccan 206
srn Sirionó 176.6
sro Siroi 57.14
srr Serrano 108.6
sru Selaru 31.2
sry Siraiya 31.4
ssa Sisaala 3.5
ssh Shinassha 7.5
ssl South Korean Sign Language 207
sss Salish (Samish Straits) 110.2
sst Salish (Straits) 110.2
sta Santa 11.1
sti Stieng 29.1.2
stl Santali 29.2
stn Sotho (Northern) 3.3.1
sto Stoney 104
sub Subiya 3.3.1
sud Sudest 31.3.1
sue Suena 57.4
sug Sungor 6.3.9
sui Sui 28.2
suk Suki 67
sul Sulka 77
sum Sumu 134
sun Sundanese 31.6.10
sup Supyire 3.5
sur Sursurunga 31.3.1
sus Susu 3.10.2
sva Svan 21
svc Slovincian 9.10
svk Slovak 9.10
svs Savosavo 82
svt Svenska Teckenspråket 207
swa Swahili 3.3.1
swe Swedish 9.5
swi Sawai 31.3.2
swr Shoshone (Wind River) 108.5
swt Swati 3.3.1
swv Swedish (Västerbotten) 9.5
syg Saryg Yughur 11.3
syu Salt-Yui 57.5
tab Taba 31.3.2
tac Tacana 186
taf Taiof 31.3.1
tag Tagalog 31.6.3
tah Tahitian 31.3.1
tai Tuareg (Air) 7.1
taj Tajik 9.8
tak Takia 31.3.1
tal Talinga 3.3.1
tam Tamang 26.2.3
tan Tangale 7.2.4
tao Tarao 26.2.7
tap Taiap 54
taq Taiwanese 26.1
tar Tariana 174
tas Tashlhiyt 7.1
tat Tatana' 31.6.1
tau Tauya 57.14
taw Tawala 31.3.1
tay Tayo 206
taz Talysh (Azerbaijan) 9.8
tba Tabare 57.5
tbb Tübatulabal 108.10
tbe Tigré (Beni Amer) 7.6
tbl Tabla 42
tbo Tboli 31.6.7
tbr Tabaru 32.3
tbs Tabassaran 20.1.3
tbt Tobati 31.3.1
tbu Tubu 6.9
tbw Tabwa 3.3.1
tce Tarahumara (Central) 108.7
tda Toda 25.4
tec Teco 132
teh Tehuelche 204.1
tel Telugu 25.3
tem Tem 3.5
ten Tennet 6.3.8
teo Teop 31.3.1
tep Tepehua 126
ter Tera 7.2.1
tes Teso 6.3.5
tet Tetela 3.3.1
tew Tewa (Arizona) 105
tex Turkic (East-Central Xorasan) 11.3
tga Tangga 31.3.1
tgb Tagbanwa (Aborlan) 31.6.3
tgh Tuareg (Ghat) 7.1
tgk Tigak 31.3.1
tgl Tshangla 26.2.3
tgn Tugun 31.2
tgo Tsogo 3.3.1
tgp Tanglapui 57.19
tgr Tigré 7.6
tgw Tarangan (West) 31.2
tha Thai 28.2
thd Thadou 26.2.7
thk Thakali 26.2.3
thn Thangmi 26.2.3
tho Thompson 110.3
thp Thaypan 84.17
ths Thai Sign Language 207
tht Tehit 32.5
thu Thulung 26.2.3
thw Thao 31.4
thy Thaayorre 84.17
tib Tibetan (Standard Spoken) 26.2.3
tic Ticuna 160
tid Tidore 32.3
tif Tifal 57.18
tig Tigrinya 7.6
tik Tikar 3.3.1
til Tillamook 110.4
tim Timugon 31.6.1
tin Tinrin 31.3.1
tir Tiriyo 175
tis Tibetan (Shigatse) 26.2.3
tiv Tiv 3.3.1
tiw Tiwi 84.19
tja Tiipay (Jamul) 117.7
tke Tokelauan 31.3.1
tki Tuki 3.3.1
tkl Takelma 115
tkm Turkmen 11.3
tks Teke (Southern) 3.3.1
tla Tolai 31.3.1
tlb Tatar-Noghay (Alabugat) 11.3
tld Talaud 31.6.9
tlf Telefol 57.18
tli Tlingit 87.3
tll Taulil 78
tlo Tobelo 32.3
tlp Tlapanec 128
tls Talysh (Southern) 9.8
tma Tama 6.3.9
tmc Timucua 92
tmg Tamagario 58
tmi Tatar (Mishar) 11.3
tmk Tumak 7.2.2
tml Tamil 25.4
tmm Tamabo 31.3.1
tmn Temein 6.3.10
tmo Tibetan (Modern Literary) 26.2.3
tmp Tampulma 3.5
tmr Temiar 29.1.1
tmu Tat (Muslim) 9.8
tna Turkana 6.3.5
tnb Tunebo 135.2
tnc Tanacross 87.1
tnd Tindi 20.1.1
tne Temne 3.2.2
tng Tongan 31.3.1
tni Tinani 26.2.3
tnj Tanaina 87.1
tnk Tungak 31.3.1
tnl Tanana (Lower) 87.1
tnn Tunen 3.3.1
tno Tondano 31.6.9
tns Tanna (Southwest) 31.3.1
tnt Tontemboan 31.6.9
tny Tenyer 3.5
toa Toaripi 72
tob Toba 199
tod Tod 26.2.3
tof Tofa 11.3
toj Tojolabal 132
tok Tarok 3.3.9
tol Tol 133
ton Tonkawa 99
toq Toqabaqita 31.3.1
tor Toratán 31.6.9
tos Totonac (Sierra) 126
tot Totonac (Misantla) 126
tou Toussian 3.5
tow Toussian (Win) 3.5
toy Tasmanian (Oyster Bay to Pitwater) 85
toz Tonga (in Zambia) 3.3.1
tpa Totonac (Papantla) 126
tpc Tepecano 108.8
tpi Tok Pisin 206
tpn Tepehuan (Northern) 108.8
tpr Tupuri 3.1
tps Tepehuan (Southeastern) 108.8

tpt Tapieté 176.6
trb Teribe 135.8
trc Trique (Chicahuaxtla) 124.4
trd Toraja 31.6.9
trg Tewa (Rio Grande) 105
tri Trique (Copala) 124.4
trm Tirmaga 6.3.8
trn Terêna 174
trr Tairora 57.8
trt Ternate 32.3
tru Trumai 178
trw Torwali 9.7
try Tiruray 31.6.7
tsa Tsakhur 20.1.3
tse Timorese 31.2
tsf Tsafiki 140
tsg Tausug 31.6.3
tsh Tümpisa Shoshone 108.5
tsi Tsimshian (Coast) 113.6
tsj Tewa (San Juan Pueblo) 105
tsl Tanzania Sign Language 207
tsm Tasmanian 85
tsn Tsonga 3.3.1
tso Tsou 31.5
tsp Tamil (Spoken) 25.4
tsr Taushiro 161
tst Tsat 31.6.10
tsw Tswana 3.3.1
tsz Tsez 20.1.1
ttb Tatar (Baraba) 11.3
tte Tutelo 104
ttn Tetun 31.2
tts Tati (Southern) 9.8
ttu Tsova-Tush 20.2
tty Tatuyo 145
tua Tuamotuan 31.3.1
tub Tubar 108.9
tuc Tucano 145
tug Tuareg (Ahaggar) 7.1
tui Türk Isaret Dili 207
tuk Tukang Besi 31.6.9
tul Tulu 25.4
tum Tumleo 31.3.1
tun Tunica 95
tup Tupi 176.6
tur Turkish 11.3
tus Tuscarora 91.1
tut Tutchone (Northern) 87.1
tuv Tuvan 11.3
tuy Tuyuca 145
tvl Tuvaluan 31.3.1
tvo Tatar 11.3
twa Twana 110.2
twe Tarahumara (Western) 108.7
twn Tiwa (Northern) 105
tws Tiwa (Southern) 105
twx Turkic (West Xorasan) 11.3
txj Totonac (Xicotepec de Juárez) 126
tye Tyeraity 84.3.4
tza Tzeltal (Aguacatenango) 132
tzb Tzeltal (Bachajón) 132
tze Tzeltal 132
tzi Taiwanese Sign Language (Ziran Shouyu) 207
tzo Tzotzil 132
tzs Tzotzil (San Andrés) 132
tzt Tzeltal (Tenejapa) 132
tzu Tzutujil 132
tzz Tzotzil (Zinacantán) 132
uby Ubykh 19
udh Udihe 11.2
udi Udi 20.1.3
udm Udmurt 10.1.1
ugs Ugandan Sign Language 207
uhi Uradhi 84.17
ukr Ukrainian 9.10
uku Upper Kuskokwim 87.1
ulc Ulcha 11.2
uld Uldeme 7.2.1
uli Ulithian 31.3.1
uma Uma 31.6.9
umb UMbundu 3.3.1
umu Umaua 175
una Una 57.17
ung Ungarinjin 84.21
unm Unami 89.1
ura Ura 31.3.1
urd Urdu 9.7
urh Urhobo 3.3.4
uri Urim 51.2
urk Urubú-Kaapor 176.6
url Urak Lawoi' 31.6.10
urm Urum 11.3
urn Urarina 164
uru Uru 193
usa Usan 57.14
usl Urubú Sign Language 207
usr Usarufa 57.8
ute Ute 108.5
uyg Uyghur 11.3
uzb Uzbek 11.3
uzn Uzbek (Northern) 11.3
vaf Vafsi 9.8
vag Vagla 3.5
vai Vai 3.10.2
vas Vasavi 9.7
vat Vata 3.8
ved Vedda 9.7
ven Venda 3.3.1
vep Veps 10.1.1
vie Vietnamese 29.1.9
vla Vlaamse Gebarentaal 207
vot Votic 10.1.1
wag Wagiman 84.7.8
wah Wahgi 57.5
wai Wai Wai 175
wak Wakhi 9.8
wal Walman 51.3
wam Wambaya 84.20
wan Wangkumara 84.17
wao Waorani 157
wap Wappo 120.1
war Wari' 182
was Washo 119
wat Watjarri 84.17
wau Waunana 136
way Wayampi 176.6
wbn Wambon 57.3
wch Wichí 198
wdo Western Desert (Ooldea) 84.17
wed Wedau 31.3.1
wel Welsh 9.4
wem Wembawemba 84.17
wer Weri 57.11
wet Wetan 31.2
wgl Waigali 9.7
wgu Warrungu 84.17
wic Wichita 103
wik Wikchamni 113.9
win Wintu 113.8
wir Wirangu 84.17
wiy Wiyot 89.2
wlf Wolof 3.2.1
wll Wallisian 31.3.1
wlm Walmatjari 84.17
wlo Wolio 31.6.9
wlw Warluwara 84.17
wly Wolaytta 7.5
wma West Makian 32.3
wmn Wéménugbé 3.9
wmu Wik Munkan 84.17
wn Wik Ngathana 84.17
wnb Winnebago 104
wnt Wantoat 57.10
wob Wobe 3.8
wog Wogamusin 52.5
woi Woisika 57.19
wol Woleaian 31.3.1
wom Womo 45.1
wor Worora 84.21
wps Wapishana 174
wra Warao 153
wrb Warrnambool 84.17
wrd Wardaman 84.7.8
wrg Warrgamay 84.17
wrk Warekena 174
wrl Warlpiri 84.17
wrm Warembori 36
wrn Warndarang 84.13
wrp Waropen 31.3.2
wrs Waris 46
wrw Warrwa 84.16
wry Waray (in Australia) 84.7.7
wsk Waskia 57.14
wth Wathawurrung 84.17
wtm Watam 53.3
wuc Wu (Changzhou) 26.1
wur Waurá 174
wwr Woiwurrung 84.17
wwy Waray-Waray 31.6.3
wya Wyandot 91.1
wyn Wayana 175
xam /Xam 1.4
xar Xârâcùù 31.3.1
xas Xasonga 3.10.2
xav Xaváante 177.2
xer Xerénte 177.2
xho Xhosa 3.3.1
xia Xiamen 26.1
xok Xokleng 177.2
xoo !Xóõ 1.4
xun !Xun (Ekoka) 1.2
yag Yagua 159
yah Yahgan 205
yak Yaka 3.3.1
yal Yale (Kosarek) 57.17
yam Yaminahua 168
yan Yana 117.6
yao Yao (in Malawi) 3.3.1
yap Yapese 31.6.11
yaq Yaqui 108.2
yar Yareba 73
yav Yavapai 117.7
yaw Yawa 35
yay Yay 28.2
yaz Yazgulyam 9.8
yba Yamba 3.3.1
ych Yup'ik (Chevak) 86
ycn Yucuna 174
yct Yucatec 132
ydb Yiddish (Bessarabian) 9.5
ydd Yiddish 9.5
ydl Yiddish (Lodz) 9.5
yei Yei 63
yel Yelî Dnye 75
yem Yemba 3.3.1
yes Yessan-Mayo 52.4
yey Yeyi 3.3.1
ygd Yag Dii 3.1
ygn Yaghnobi 9.8
ygr Yagaria 57.8
yi Yi 26.2.4
yid Yidiny 84.17
yil Yil 51.3
yim Yimas 53.4
yin Yindjibarndi 84.17
yir Yir Yiront 84.17
yki Yuki 120.2
ykm Yakoma 3.1
ykn Yakan 31.6.6
yko Yukaghir (Kolyma) 12
ykp Yukpa 175
ykt Yakut 11.3
ylb Yulparija 84.17
yli Yali 57.6
ylm Yelmek 62
ylt Yiddish (Lithuanian) 9.5
ymd Yamdena 31.2
ymi Yami 31.6.4
yms Yemsa 7.5
yna Yupik (Naukan) 86
yng Yingkarta 84.17
ynk Yankuntjatjara 84.17
ynm Yanomámi 154
yns Yansi 3.3.1
yny Yanyuwa 84.17
yok Yokuts (Yaudanchi) 113.9
yor Yoruba 3.3.3
ypk Yup'ik (Central) 86
yqy Yaqay 59.1
yrc Yuracare 191
yrm Yurimangí 138
yrr Yaruro 152
yrt Yuruti 145
ysi Yupik (Sirenik) 86

ysl Yupik (St. Lawrence Island) 86
yta Yurt Tatar 11.3
ytu Yukaghir (Tundra) 12
yuc Yuchi 93
yug Yugh 13
yuk Yukulta 84.18
yul Yulu 6.2.1
yun Yup'ik (Norton Sound) 86
yur Yurok 89.3
yus Yupik (Siberian) 86
yuw Yuwaalaraay 84.17
ywl Yawelmani 113.9
ywr Yawuru 84.16
yyg Yaygir 84.17
yyo Yorta Yorta 84.17
yzv Yazva 10.1.1
zai Zapotec (Isthmus) 124.8
zaj Zapotec (Juárez) 124.8
zam Zapotec (Mixtepec) 124.8
zan Zande 3.1
zap Zapotec (Mitla) 124.8
zaq Zapotec (Quiegolani) 124.8
zar Zarma 6.11
zay Zayse 7.5
zaz Zazaki 9.8
zch Zoque (Chimalapa) 130
zen Zenaga 7.1
zfl Zoque (Francisco León) 130
zha Zhang-Zhung 26.2.3
zhn Zhuang (Northern) 28.2
zim Zimakani 59.1
zno Zulu (Northern) 3.3.1
zpi Zapotec (Ixtlan) 124.8
zpr Zaparo 158
zqc Zoque (Copainalá) 130
zqo Zoque (Ostuacan) 130
zqr Zoque (Rayon) 130
zqs Zoque (Soteapan) 130
zso Zulu (Southern) 3.3.1
zsq Zapotec (San Lucas Quiaviní) 124.8
zte Zapotec (Texmelucan) 124.8
zul Zulu 3.3.1
zun Zuni 107
zya Zapotec (Yatzachi) 124.8
zzo Zapotec (Zoogocho) 124.8